Shame As It Ever Was: Stigma and Personal Bankruptcy

Kartik Athreya

Bankruptcy rates have risen rapidly over the past two decades, most noticeably since approximately 1991, with rates doubling by 1997. To account for this rise in bankruptcy rates, researchers have advanced two—perhaps competing—hypotheses. The first hypothesis is that the availability of unsecured consumer credit has expanded to wider, riskier subsets of the population in response to cost reductions arising from improvements in technology for monitoring and accounting for borrowers' activities. Further, technological improvements in payment networks have promoted the widespread adoption of credit cards as a method of payment. The recent works of Ellis (1998), Athreya (2001), and Moss and Johnson (2002) illustrate the phenomenon of rising credit card usage as it relates to increases in bankruptcy.

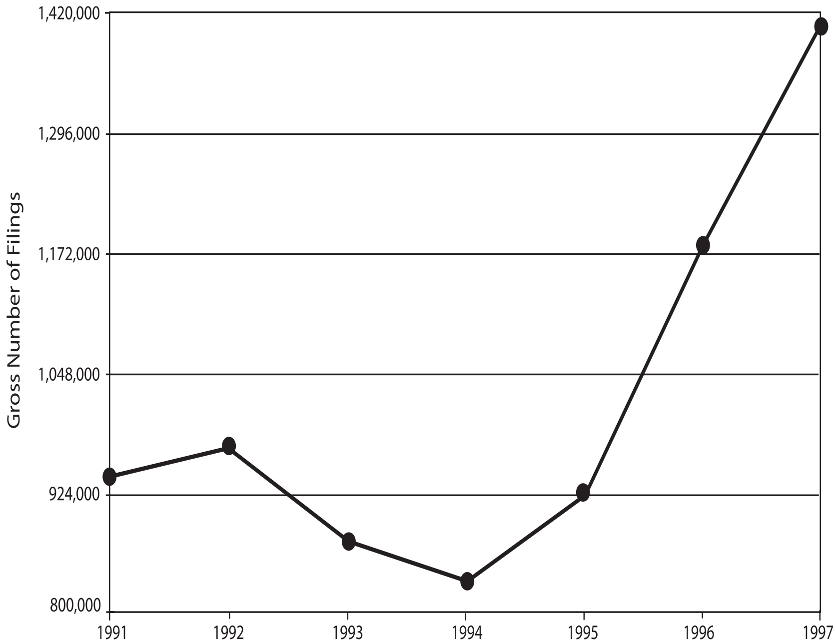
A second view, and one that appears commonly held among the general public, is that bankruptcy has lost its once formidable stigma. Gross and Souleles (2002) argue that non-risk-related factors have become more important in explaining the rise in bankruptcies and attribute this to a fall in stigma. Fay, Hurst, and White (forthcoming) propose that stigma was higher in the past and that bankruptcy filing rates were therefore lower. For this reason, current research is divided on the cause of the rise in bankruptcy rates.

Anecdotally, it is not difficult to find reports like the following:

> Judge Edith Jones, a former member of the National Bankruptcy Review Commission, noted in Congressional testimony that "The current system

I thank Tom Humphrey, Yongsung Chang, John Walter, and John Weinberg for very helpful comments. The views expressed are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System. All errors are mine.

Figure 1 Total Bankruptcy Filings, United States, 1991–1997



> of bankruptcy law permits any person to seek relief without demonstrating financial necessity. At one time in our history, filing bankruptcy was regarded as shameful, and filers suffered social stigma and permanently ruined credit. The shame and stigma are no longer compelling. . . many filers now commence cases without ever having been in default on their debts. This suggests that bankruptcy is, to them, not a last resort, but a first resort" (Pomykala 1999).

A recent article in *Harvard Magazine* (Anderson 1999) also implies that stigma has fallen by noting that "Industry representatives point to the fact that personal bankruptcy filings hit an all-time high of 1.4 million in 1997, during a period of low unemployment, when consumers were in their best financial shape in many years."[1]

The purpose of this article is twofold. First, the work of the preceding authors is empirical, not "structural" in the sense that households are not modeled explicitly. An innovation of this article is to provide a clear exposition of a

[1] This statement is misleading in the sense that while the average household experienced gains during the 1990s, it is also the case that those in bankruptcy were typically in very difficult circumstances (Sullivan et al. 2000).

structural model of lending and borrowing with bankruptcy as well as attempt to account endogenously for the expansion in unsecured credit availability and bankruptcy. This article therefore goes beyond that of Athreya (2001), where the expansion of credit was taken as exogenously given. This analysis is meant to illustrate how bankruptcy affects demand and supply in the unsecured credit market. The second goal of this article is to employ the model to study the observable implications of the widely held belief that bankruptcy is far less painful today than it was in the past.

A central challenge for explanations of rising bankruptcy rates based on falling stigma is that such a feature should generate a supply-side response whereby borrowing on the unsecured credit market grows more expensive. That is, if borrowers become more willing to default, lenders will seek higher interest rates to compensate for the greater default risk. In turn, such a price increase should be associated with smaller debt holdings across households, something that is counterfactual for the period under study. I exploit precisely this implication by studying stigma in a model where creditors must at least partially respond to borrower characteristics when pricing loans.

In this article I show that decreasing the nonpecuniary cost of bankruptcy, as a fall in stigma implicitly does, indeed increases bankruptcy rates but yields counterfactual implications for the time path of debt held by households. I then address the consequences of making credit more widely "available" by reducing transactions costs for intermediaries. In particular, I show that minimally vigilant creditors and optimizing households will act in a manner that prevents a fall in stigma from generating the simultaneous rise in bankruptcy and debt holdings observed in the data. Therefore, a main conclusion of this article is that stigma is by no means dead. A more plausible explanation for the simultaneous run-up in both unsecured debt levels and bankruptcy rates may lie in the implicit reduction in costs of intermediation that occurred over the past two decades. Fully accounting for the facts, however, remains a task for future work.

1. THE FACTS

What are the main facts describing bankruptcy rates and their evolution over the past two decades? Also, have the characteristics of filers changed during the period? The data presented here draw heavily on the detailed study of bankruptcy filers of Sullivan et al. (2000).

One can immediately see that bankruptcy rates have risen somewhat steadily over time, with the most spectacular rise coming in the period from 1990 to the present. Figure 1 illustrates the period from 1990 onward.

- Fact 1: Bankruptcy rates roughly doubled during the period 1991–1997.

Figure 2 Outstanding Consumer Revolving Debt as a Percentage of Personal Disposable Income in the United States, 1991–1997

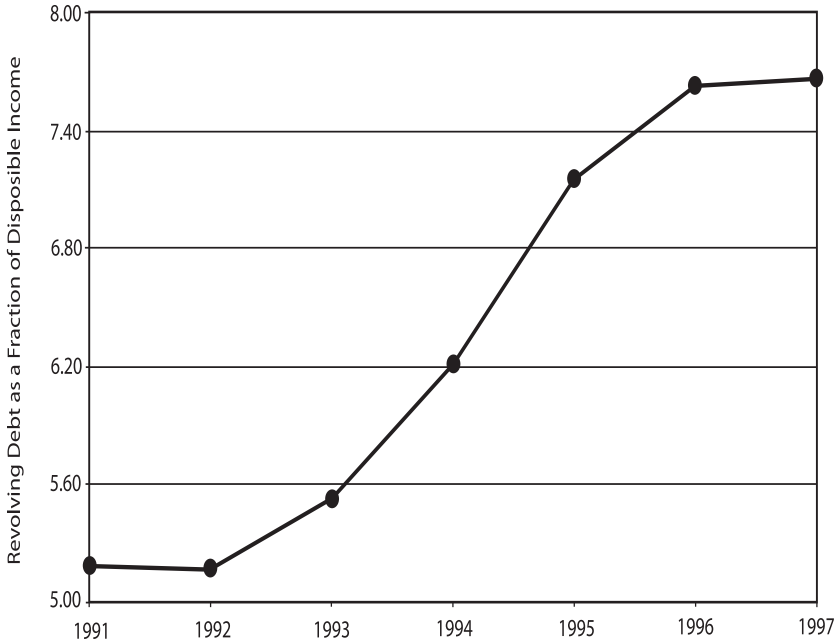


Figure 2 documents a measure of the indebtedness of households over the same period. The figure shows the ratio of the stock of revolving debt to personal disposable income. As is clearly visible, households did become more indebted along this dimension relative to their incomes during this period.

- Fact 2: U.S. households as a whole sharply increased their holdings of revolving debt relative to income over the period 1991–1997.

To show a more precise idea of the "profile" of a bankruptcy filer and how it has changed over time, Sullivan et al. (2000) document statistics on the characteristics of bankrupt households, presented in Table 1.

Notable in this table is the rise in the ratio of median unsecured debt in bankruptcy relative to median income at the time of filing.[2] In 1991, this ratio

[2] All debt and income levels are in 1997 dollars.

Table 1 Characteristics of Bankruptcy Filers in 1991 and 1997 (in 1997 dollars)

1991	**Income**	**Unsecured Debt**
Mean	23,927	26,618
Standard deviation	14,357	48,461
25th percentile	14,141	8,208
Median	21,155	15,128
75th percentile	31,110	26,934
N	641	687
Missing	95	49

1997	**Income**	**Unsecured Debt**
Mean	19,641	29,529
Standard deviation	10,482	37,618
25th percentile	12,072	12,368
Median	18,756	19,515
75th percentile	25,748	31,389
N	100	100
Missing	0	0

was roughly 0.75 (i.e., $15,128/$21,155). By 1997, the ratio had exceeded unity, at $19,515/$18,756.[3] Therefore, we have:

- Fact 3: Debt-to-income ratios for U.S. households in bankruptcy increased sharply over the period 1991–1997.

Despite the fall in median income at the time of filing during the period 1991–1997, one should note that median household income for the entire U.S. population remained essentially fixed, at roughly $40,000 (in current dollars). While the income process in the model is too discrete to allow an exact comparison with these statistics, it still will be possible to provide an approximation of these statistics to help evaluate the predictions of the falling stigma.

I now develop a simple model of unsecured borrowing and ask if falling stigma is consistent with the facts documented above. In order for a decline in stigma to be regarded as a serious explanation for the increase in bankruptcy rates between 1991 and 1997, reductions in stigma must also generate a) sharp increases in the level of unsecured debt held by households, as well as b) large increases in the median debt discharged per bankruptcy.

[3] By contrast, in 1981, the ratio was slightly less than 0.5, as the median household in bankruptcy held $12,452 dollars of unsecured debt, relative to a median income of $26,439.

2. A MODEL OF CONSUMER DEBT AND BANKRUPTCY

The model developed here follows Athreya and Simpson (2003), and is an extension of Athreya (2002). It is richer than the latter in that creditors are allowed to condition the interest rates they charge according to an approximation of the default risk posed by each loan. Conditioning allows us to incorporate non-trivial supply side effects on credit availability, which, in turn, has implications for long-run indebtedness and default rates.

The environment under study consists of a large number of initially identical households that each receive shocks to their income in every period. Households cannot perfectly insure their income risk. Instead, they have access only to risk-free savings and debt that may be discharged subject to penalties associated with bankruptcy.

More precisely, households maximize:

$$U = E_0 \sum_{t=0}^{\infty} \beta^t u(c_t).$$

Where c_t denotes consumption within a period, β defines the rate at which households discount future consumption. In this specification, $u(.)$ denotes the period utility function, which we take to have the constant relative risk aversion (CRRA) form:

$$u(c_t) = \begin{cases} \frac{c_t^{1-\gamma}-1}{1-\gamma}, \text{ for } \gamma > 1 \\ \ln c_t, \text{ for } \gamma = 1. \end{cases}$$

The parameter γ governs how risk-averse households are, as well as how smooth they would want consumption to be over time, given prices. In each period, households draw a random endowment of income, denoted e. The randomness of income is meant to capture both variations in income among working people, such as changes in overtime, short-term disability, etc., as well as variations in income arising from becoming unemployed when working, or re-employed when not. Employed households may receive a shock that renders them unemployed. The probability of getting this shock is given by ρ. Once unemployed, the household regains employment in each period with probability π_{emp}.

Given their preferences, households would like to smooth consumption, but are assumed not to have access to explicit insurance contracts against their income risk. Instead, they must smooth consumption by either accumulating one-period risk-free claims to consumption or by borrowing. When borrowing, however, households are endowed with the option to default, which we will take as being representative of the formal process of bankruptcy. Before turning to the definition of bankruptcy, note that in this environment different debt levels will be associated with different default risks. Therefore, if financial intermediaries are allowed to observe the total debt issued by a household, competition will force firms to condition on total debt.

Let $r(a')$ denote the net interest rate associated with the amount of savings or debt a'. Households save by choosing $a' > 0$, and borrow by choosing $a' < 0$. Given that there is no default risk for savings, the interest earned on savings will not vary with the amount saved and will therefore be simply a single risk-free rate, which I denote by $r^d > 0$. That is, $r(a') = r^d$ when $a' > 0$. On the other hand, households may choose to borrow by issuing debt that matures next period. However, a central feature of the model is that agents may default on debt. I therefore follow Livshits, MacGee, and Tertilt (2003) and model the market for debt as follows.

First, assume competitive financial intermediation whereby the savings of households are taken in and then loaned to large numbers of households. Assume further that making loans entails transactions costs arising from the costs of processing applications, verification of information, etc. Let the transactions cost be denoted by τ. In this case, because of the possibility of default, the debt is effectively discounted so that $r(a') \geq r^d + \tau$. In a given period, a household must choose the amount of one-period borrowing it desires and then issue a given face-value of debt accordingly. This debt is then discounted at least at the risk-free rate. Therefore, if the household issues a given face value of debt a' in this period, it receives a smaller amount of consumption goods $\frac{a'}{1+r(a')} < a'$. In the following period, the debt matures, and the household decides either to honor the debt or to file for bankruptcy to be released of its obligation. For large debt levels, where default is relatively more likely, $r(a')$ will of course be larger.

In the limit, the household will receive nothing at all if it attempts to issue debt levels where default in the next period becomes certain, as $r(a')$ will approach infinity. This is precisely the sense in which the borrowing limits faced by households are endogenous, depending on the likelihood that the household defaults on a given amount of debt. I denote the endogenous credit limit for solvent households by $\underline{a}_s$.

Upon filing for bankruptcy, households have their debts discharged and pay two costs. First, they are restricted from future borrowing for an uncertain length of time. Specifically, in each subsequent period, the household has its ability to borrow restricted with probability $(1 - \psi) < 1$, resulting in an average length of restricted borrowing of $1/(1-\psi)$ periods. A second cost that households must pay is the stigma of bankruptcy, represented by an amount λ subtracted from the household's utility in the period it files.

The nature of the problem faced by households is such that if households derive a given value V from behaving optimally from next period onward, then their choices today are restricted by only a few variables. The first variable is household credit status. In any period, households may be solvent, S, may have just filed for bankruptcy, B, or may be "borrowing constrained" due to a past bankruptcy, denoted BC. I denote credit status by the variable $CS \in \{S, B, BC\}$. Second, households must know their current period earnings, e. Third, households must know their wealth or debt position, a. The value

V, therefore, is a function of these three pieces of information, which, when taken together, fully determine the resources available for consumption today. I therefore denote the "value function" by $V(CS, e, a)$. It is then sufficient for the household to maximize utility today, taking as given that it will choose optimally thereafter, obtaining value $V(CS', e', a')$.

Adapting slightly the notation used in Athreya (2002), we have a simple representation of the problem. The value of being solvent $V(S, e, a)$, is expressed as the better of two choices that a solvent household has in a period, namely, whether to file for bankruptcy or not. Therefore, $V(S, e, a)$ must satisfy:

$$V(S, e, a) = \max[W(S, e, a), W(B, e, a)] \tag{1}$$

where $W(S, e, a)$ denotes the value of *not* filing for bankruptcy in the current period. If the household does not file for bankruptcy, it simply chooses consumption, c, and either saves or issues debt, a', and moves to the next period as a solvent household.

Therefore, $W(S, e, a)$ satisfies:

$$W(S, e, a) = \max\{u(c) + \beta EV(S, e', a')\} \tag{2}$$

s.t.

$$c + \frac{a'}{1 + r(a')} \leq e + a \tag{3}$$

s.t.

$$a' \in A_S \tag{4}$$

where $A_S = [\underline{a}_s, \infty)$, implying that a household faces an absolute credit limit of $\underline{a}_s$.

If, on the other hand, a household does choose to file for bankruptcy, its debts are erased, as seen below in the modified right-hand side of the household budget constraint, (6), leaving only current earnings for consumption and savings. The household is also hit by the stigma of filing, λ. It is λ that will be calibrated in the benchmark and then altered. It is the central parameter in the analysis that follows. In the following period, the credit market is assumed to shut the household out, and the latter obtains a value $V(BC, e, a)$ from beginning in this state. The value of filing for bankruptcy in a given period is then given by:

$$W(B, e, a) = \max\{u(c) - \lambda + \beta EV(BC, e', a')\} \tag{5}$$

s.t.

$$c + \frac{a'}{1 + r^d} \leq e \tag{6}$$

s.t.

$$a' \in A_{BC.} \tag{7}$$

If a household files for bankruptcy, it is constrained from borrowing in the following period. In subsequent periods, the restriction on borrowing ends with a time-independent probability ψ. The set $A_{BC} = [\underline{a}_{BC}, \infty)$, where $0 \geq \underline{a}_{BC} \geq \underline{a}_S$. Given the preceding, the maximal utility attainable by a borrowing-constrained household satisfies equation (8) below, subject to the constraints in (9) and (10).

$$V(BC, e, a) = \max\{u(c) + \psi\beta EV(S, e', a') + (1-\psi)\beta EV(BC, e', a')\} \tag{8}$$

s.t.

$$c + \frac{a'}{1+r^d} \leq e + a \tag{9}$$

s.t.

$$a' \in A_{BC}. \tag{10}$$

3. EQUILIBRIUM

It can be shown that the savings and debt instruments allowed here are in general insufficient for households to fully protect their consumption of goods and services from fluctuations in their income. Therefore, as time evolves, households will hold different wealth levels as a function of their past incomes. Under fairly general conditions, however, the distribution of wealth across households will converge to a distribution that, once reached, will remain fixed. I will focus on the allocations associated with this fixed "long-run" distribution.

Equilibrium further requires that households act optimally; that is, their decisions must satisfy the recursive representation above, and lenders must break even. With respect to the latter, a loan of size a carries an associated default probability $\theta^{bk}(a)$ and a cost of funds (including transactions costs) of $(r^d + \tau)$. Therefore, the zero-profit interest rate for loans $r(a)$, must satisfy:

$$r(a) = \frac{(r^d + \tau)}{(1 - \theta^{bk}(a))}. \tag{11}$$

For those who choose to hold savings, the zero-profit interest rate will simply be r^d.[4]

4. THE EXPERIMENT

The experiment I conduct is very simple. I first set model parameters, chief among them stigma, such that the model approximately matches bankruptcy

[4] I refer the interested reader to Athreya (2002) for more details on the stationary equilibrium concept employed here.

Table 2 The Change in Bankruptcy and Credit: 1991 and 1997

	1991	**1997**
Bankruptcy rate	0.13%	0.20%
Median debt/income in bankruptcy	38%	50%
Median debt in bankruptcy	\$15,100	\$19,500
Revolving debt/income	5.5%	8.5%
Credit card charge-off rate	0.85%	1.25%

rates, the debt discharged in bankruptcy, and charge-off rates as of 1991. I then lower the cost of stigma and ask what the model predicts for bankruptcy rates, debts discharged in bankruptcy, and interest rates. If the model is consistent with observed outcomes in 1997, it will provide some support for the idea of lower bankruptcy stigma in the present than in the past. If not, we will be left with the puzzle of why the observed changes have occurred. The observed outcomes as of 1991 and 1997 are given in Table 2. I calculate the bankruptcy target as follows. In 1991, there were approximately 900,000 filings. Of these, 70 percent were Chapter 7 filings, to which the bankruptcy procedure defined in the model best matches. Further, Sullivan et al. (2000) report that roughly 20 percent of bankruptcies involved health problems, and I have excluded those cases, as they are better modeled as shocks to "expenses," as in Livshits et al. (2003). I therefore target an annual filing incidence of 500,000. Given the approximately 100 million households in the United States, I obtain a quarterly filing rate of 0.13 percent. For the ratio of revolving debt to income, I use data from the *Economic Report of the President* (2001). Charge-off rates are taken from Stavins (2000), and the remaining facts are from Sullivan et al. (2000).

Parameters

The exogenous process for labor income shocks is given below and will take four different values: $Y \in \{Y_1, Y_2, Y_3, Y_4\}$. These shocks represent a proportional increment or decrement to average productivity. The shocks are set to allow the interpretation that households are partitioned into those who are "working," and those who are "unemployed." When households are working, income shocks take values in the set $Y_w = \{Y_1 = 1.15, Y_2 = 0.85\}$, with probabilities that depend on the previous period's shock. In each period, an employed household may receive a shock that renders it unemployed. The probability of this shock is denoted by ρ. When they are unemployed, households will draw income from a set of two substantially lower shocks, Y_3 and Y_4, where $Y_3 > Y_4$, and Y_j is a positive income amount for all j. The size of these shocks is set at: $Y_u = \{Y_3 = 0.5, Y_4 = 0.1\}$. These shock realizations

Table 3 The Income Process

Employed	$Y' = Y_1$	$Y' = Y_2$	$Y' = Y_3$	$Y' = Y_4$
$Y = Y_1$	0.9700	0.0300	0.0000	0.0000
$Y = Y_2$	0.0300	0.9700	0.0000	0.0000
Unemployed $<$ 2 Quarters	$Y' = Y_1$	$Y' = Y_2$	$Y' = Y_3$	$Y' = Y_4$
$Y = Y_3 \mid high\ productivity$	0.6790	0.0210	0.1800	0.1200
$Y = Y_4 \mid low\ productivity$	0.0210	0.6790	0.1800	0.1200
Unemployed $\geq$ 2 Quarters				
$Y = Y_3 \mid high\ productivity$	0.6790	0.0210	0.0000	0.3000
$Y = Y_4 \mid low\ productivity$	0.0210	0.6790	0.0000	0.3000

are interpreted as income received by those who qualify and those who fail to qualify for unemployment insurance. To avoid keeping track of a separate shock process for unemployed households, it is assumed that they receive high and low productivity draws even when unemployed from the same process as employed households. Therefore, if re-employed in the following period, these households will receive incomes according to the conditional probabilities of low and high productivity shocks given current period productivity. Lastly, households hit by two or more consecutive periods of unemployment no longer qualify for unemployment insurance and instead receive the shock Y_4. Table 3 describes the transition probabilities.

For expositional ease, I display the earnings process for households, given that the employment/unemployment shock has already been realized. The interpretation of this process is the following: For employed workers, the process is given by the top two rows of Table 3. Since they are employed in the current period, they face zero probabilities of drawing income levels Y_3 and Y_4. Instead, their productivity remains within the set $\{Y_1, Y_2\}$ and is drawn according to the corresponding transition probabilities. Next, consider the process for a newly unemployed worker with currently high individual productivity. This worker has a probability of 0.679 of becoming employed with high productivity and a probability of 0.0210 of becoming employed with low productivity. Additionally, the worker will remain unemployed with a probability of 0.30 (i.e., 0.18 + 0.12) and will retain high productivity with probability 0.18 and loses productivity with probability 0.12. Finally, consider a worker who is unemployed for at least two periods, as shown in the bottom two rows of Table 3. Such workers will no longer qualify for unemployment insurance benefits should they fail to find work, which occurs with probability 0.30 in a given quarter. Therefore, these households draw incomes of $Y = Y_4$. If the household does find work, it then draws from the process $Y_w = \{Y_1 = 1.15, Y_2 = 0.85\}$.

Table 4 The Parameters

Parameter	Description	Value
γ	Risk aversion	1.00
τ	Quarterly transition cost of intermediation	0.0085
ρ	Unemployment risk (quarterly)	0.05
π_{emp}	Probability of re-employment in next period	0.7
$1/(1-\psi)$	Avg. length of post-bankruptcy borrowing limit (qtrs)	$16(\Rightarrow \psi = 0.9375)$
λ	Bankruptcy cost	1.65*
β	Quarterly discount rate	0.9865*
$\underline{a}^B$, $\underline{a}^{BC}$	Borrowing limit for borrowing-constrained households	0

Table 4 below reports the remaining parameters of the model that, along with the income process, matches outcomes as of 1991. Asterisks indicate calibrated parameters. That is, parameters whose values are set such that the model matches the facts documented above, as of 1991. In 1991, the unemployment rate was 6.8 percent. The fact that 1991 was the trough of a business cycle implies that unemployment risk at this time was higher than in 1997. As a result, the benchmark estimate of stigma might be affected by the value imputed for the risk of becoming unemployed, ρ. However, the direction of any bias in measuring stigma under 1991 unemployment rates is unclear, as there are forces in both directions. Namely, in a world with relatively high unemployment risk, bankruptcy is more costly than when unemployment risk is low, as exclusion from credit markets becomes more painful. This will lead to lower debt holdings for households, all else equal. Conversely, more unemployment risk means that those with large debts are likely to find themselves in more difficult circumstances than before. It should be noted that bias in estimation of stigma arising from a higher or lower unemployment target may arise. Such bias, however, is likely to be second order with respect to the qualitative changes induced by a large fall in stigma.

The remaining parameters are chosen either with direct reference to the data, as in the case of τ and ψ, or are set at values made standard in the literature, as in the case of the risk-aversion coefficient γ and the discount factor β.[5] The stigma-related cost of bankruptcy λ, is a reduction in utility arising from filing. To be concrete, it is useful to express this reduction in terms of the consumption good or in dollars. To do this, note that the value function yields important information on the consequences of a reduction in current "cash-in-hand" for a household contemplating bankruptcy. Namely, we can ask the following questions: First, what are the financial characteristics (i.e., current earnings, e) of the median bankruptcy filer in our model? For this

[5] For the first two parameters, see Athreya (2002) for details.

Table 5 The Benchmark Calibration

	1991 Data	Model
U.S. quarterly bankruptcy rate	0.13%	0.14%
Median debt/income in bankruptcy	38%	34%
Median debt in bankruptcy	$15,000	$13,700
Revolving debt/income	5.2%	6.2%
Unemployment rate	6.8%	6.9%
Credit card charge-off rate	1%	0.76%

value of earnings, how much would current earnings, e, have to be reduced in order for an optimizing household to find its utility lowered by λ units?[6] This value turns out to be approximately $4,000. Therefore, bankruptcy is not taken lightly by households. This is also consistent with the work of White (1998) who finds that while 15 percent of households would find it financially beneficial, less than 1 percent do file in any given year.

5. RESULTS

In the benchmark case, we see that the model matches up quite well with the data along several dimensions, even though there are more targets than "free" parameters. In particular, with respect to credit market facts, I attempt to match the data along three dimensions: 1) the U.S. quarterly bankruptcy rate, 2) the ratio of median debt to median income among bankruptcy filers as well as the dollar value of the debt, 3) the ratio of total revolving debt to income for the United States, 4) the unemployment rate, and 5) the quarterly loss-rate or "charge-off" rate for credit cards—the predominant form of unsecured debt.

[6] More precisely, we compute the following: let $x = \{CS, a, e\}$ be any value for the state vector, and let e_{med} be the median value for earnings among those filing for bankruptcy. Given our notation, the expected discounted utility from filing for bankruptcy is then given by:

$$W(B, e_{med}, a) = \max\{u(c) - \lambda + \beta EV(BC, e', a')\} \tag{12}$$

s.t.

$$c + \frac{a'}{1 + r^d} \leq e_{med} \tag{13}$$

s.t.

$$a' \geq \underline{a}_{BC} \tag{14}$$

With this, we can solve for the reduction in income such that optimizing under the remaining income, e^*, would yield utility level, $W(B, e_{med}, a)$. This will tell us how costly in real terms the penalty is to a representative member of the group who receives penality costs. In the model, the median unemployed household earns a median income of approximately 0.5 units. The median filer, on the other hand, receives an income of 0.1 units. This difference of 0.4 units of current earnings produces a utility drop of roughly 1.5 units. In dollar terms, 0.4 is 40 percent of quarterly earnings or $4,000.

Given that the benchmark model captures many of the salient features of bankruptcy, unemployment, and credit, I now evaluate the effect of a reduction in "stigma," denoted by the parameter λ. In order to avoid further lowering bankruptcy by lowering the unemployment rate, I leave the probability of unemployment, ρ, fixed, even though 1997 was characterized by a lower unemployment rate than 1991. If the model is still unable to generate an increase in bankruptcy as stigma falls, this will provide a more compelling argument against the stigma-based explanation for the observed changes. In the results above, λ is cut from its 1991 benchmark value of 1.65 to a value of 0. Beyond exclusion from credit markets, there is assumed to be no cost to bankruptcy whatsoever. This change in stigma, while perhaps stark, is useful for two reasons. The complete elimination of stigma may help generate an increase in the bankruptcy rate consistent with the enormous increase observed in the data. Additionally, in this model, I have found that the response of stigma is qualitatively similar for smaller changes, and therefore, there is no great loss in generality by considering the limiting case. I eliminated exclusion from credit markets as part of "stigma" because the common interpretation of stigma involves personal nonpecuniary suffering associated with the decision to file for bankruptcy.

It is seen that bankruptcy rates do rise, almost to levels observed in 1997. Therefore, the hypothesis that there has been a loss in stigma is not yet contradicted. However, what has happened to credit "supply" relative to the benchmark is striking. Table 6 shows that the endogenous borrowing limit, "$\underline{a}_s$," is only approximately \$1,100, or *less than three percent* of annual median income. In essence, the unsecured credit market nearly disappears. By contrast, when the stigma of bankruptcy is set at the 1991 benchmark level, the endogenous limit on unsecured debt rises dramatically to the equivalent of one-third of annual income, or roughly \$14,000. The debt discharged in bankruptcy under low-stigma falls to nearly zero. In this low-stigma environment, households rarely borrow, but when they do, they often default. The latter is seen in the quarterly charge-off rate, which at nearly 13 percent is wildly at odds with the roughly 1 percent rate observed in the period since approximately 1996. The results for credit loan rates are predicated on a creditor who can observe total debt. To the extent that lenders could observe more, one might expect the effects on supply to be even more severe than reported here.

Intuition

To help explain the decisionmaking of households, it is important to understand households' willingness to acquire debt and, in turn, pay premia associated both with transactions costs and with default risk. In particular, the benefits of asset accumulation and decumulation to smooth consumption are significantly reduced when shocks are very persistent—or when they are very large in magnitude. For environments similar to the present one, but without default,

Table 6 The Effect of a Fall in Stigma

	1991	1997
λ	1.65	0.00
Bankruptcy rate	0.13%	0.18%
Median debt/income in bankruptcy	34%	0.85%
Median debt in bankruptcy	$13,700	$340
Revolving debt/income	6.2%	0.2%
Credit card charge-off rate	0.8%	12.9%
$\underline{a}_s$	$14,000	$1,100

households can typically do a great deal of consumption smoothing without ever borrowing. For example, the classic work of Aiyagari (1994) demonstrates that for even very persistent (but not permanent) shocks, the additional amount of assets households accumulate relative to the full-insurance case is small. This finding is important for understanding why a reduction in the cost of bankruptcy leads, in equilibrium, to a substantially lower availability of credit. I now discuss the decision rules of households with respect to debt in the two environments.

Intuitively, in order for a reduction in stigma to generate increased bankruptcy filings, the "supply" effect in terms of the loan-rate schedule must be attenuated so that households are still able and willing to obtain credit cheaply and to discharge these debts at relatively higher rates than with higher stigma. For households to be willing to pay large premia to obtain credit, they must be likely to default or in dire circumstances, or both. However, for empirically plausible specifications of income risk, households in the model can effectively smooth consumption even without access to credit markets. Therefore, if a household in such an environment were given the same amount of credit as before (i.e., as much as in a "high stigma" world), they would now be far more likely to default on the debt. In turn, the competitive price for such a loan would rise substantially. If adequate consumption smoothing is then possible without recourse to borrowing, most households will be unwilling to pay large premia to borrow. Therefore, as stigma falls, equilibrium debt holding falls drastically even while bankruptcy rates rise. The central point therefore is that the "supply" side response of credit to costs of bankruptcy is intimately connected to how well consumers may smooth consumption without borrowing.

The central question now becomes: since falling stigma appears entirely unequal to the task of explaining the facts, is there a plausible alternative explanation? Improvements in technology for intermediating credit is one prominent possibility, which I now explore.

What About Competition and Technological Change in Unsecured Lending?

In this section, I evaluate the possibility that competitiveness in unsecured lending intensified, most notably in credit cards, while advances in computing and telecommunications further lowered the costs of maintaining an unsecured credit network, leading borrowing to rise and prices to remain steady even with a fall in stigma. Examples include the PS2000 payments authorization system introduced by VISA in 1992 and, more visibly, the vast array of co-branded cards offering frequent flyer miles and other rebates that effectively lower the price of unsecured credit. These reductions in the marginal costs of credit supply are also consistent with the substantial entry to credit card lending that took place in the period 1991–1995 and, in turn, greatly increased the number and variety of credit card contracts offered, as well as the debt taken on by households.[7] Edelberg (2003) documents very clearly that the extent to which creditors now are able to engage in risk-based pricing has increased. In particular, she finds a large increase in the variance in interest rates charged to different consumers during the mid-1990s, precisely the window under study in this article. The reduction in the marginal costs of account management thus allowed better partitioning of high-risk borrowers from low-risk borrowers and is also likely to have increased the lender's overall willingness to extend credit to all households. This is the heart of the argument that the 1990s were a period of "democratization" of credit. This means, by definition, that borrowing costs have fallen for a large population, which is what leads them to borrowing on mainstream unsecured credit markets, as opposed to either being denied credit or borrowing from more expensive alternatives such as check-cashing outlets or pawn shops. As Edelberg (2003) argues, in the current environment, "we expect to see that very high risk borrowers are able to get credit . . . instead of simply being denied."

The parameter τ represents transactions costs, broadly defined. These costs may be associated with both innovations in the technology of credit services and more competitive lending.[8] The natural experiment is therefore to lower τ, (set to 0.0085). To be extreme, I now evaluate outcomes for the case where $\tau = 0$. The results are presented in the second line below. In the second line of Table 7, I present the results that isolate the impact of technological change in intermediation alone, whereby I reinstate the benchmark non-pecuniary cost of bankruptcy of $\lambda = 1.65$, but drop the transactions cost τ to zero.

---

[7] See Evans and Schmalensee (1999).

[8] To the extent that lending became more competitive post-1991, the purely competitive zero-profit condition in the 1991 benchmark is not quite correct. However, it is likely to be a useful approximation and allows us to reduce τ and interpret this meaningfully as a reduction in margins.

Table 7 The Effect of Lower Transactions Costs ($\tau = 0$)

	1991	$\tau = 0$	$\tau = 0$
λ	1.65	1.65	0.00
Bankruptcy rate	0.14%	0.19% (0.20%)	0.20%
Median debt/income in bankruptcy	34.5%	40.3% (50%)	3.3%
Median debt in bankruptcy	$13,700	$16,100 ($19,000)	$1,200
Revolving debt/income	5.2%	14.9% (8.5%)	0.001%
Credit card charge-off rate	0.8%	0.5% (1.25%)	15%

For comparison, the actual 1997 data are presented in parentheses. It is clear that the reduction in transactions costs/rents produces model outcomes closer to observed changes than reductions in stigma. In the former case, bankruptcy rates and debt holdings both rise, as the data indicate, while in the latter, rising bankruptcy rates are associated with extremely low-debt levels in bankruptcy. In fact, the low transactions cost case appears to match closely the observed increase in filings, with an implied annual rate of 0.76 percent—very close to the approximately 0.80 percent rate observed in the data. This model with lower transactions costs is also able to nearly generate the $19,000 median debt per bankruptcy seen in the data. However, the model is not able to capture this increase in charge-off rates that occurred over the period from1991 to 1997, and this feature is related to the fact that households take on far more debt under the low-transactions regime than in the data. Nonetheless, the lower cost of unsecured borrowing clearly implies forces in a qualitative direction consistent with the observed changes.[9]

Conclusion

To the extent that the model presented here delivers counterfactual results when stigma is reduced, one might conclude that the model is flawed rather than interpret the results as suggesting that stigma has not dropped. However, the model is predicated on only a few plausible assumptions regarding household risk, credit markets, and incomplete insurance markets. Secondly, it was assumed that households could only imperfectly insure their labor income risk—a reasonable assumption. Thirdly, creditors are assumed to price

[9] The last column of Table 7 explores the effects of the elimination of both stigma and transactions costs. These changes result in bankruptcy rates rising only slightly relative to the benchmark but also are associated with extremely small debt levels being discharged in bankruptcy. A possible reason for the fall in debt holdings is that with the elimination of stigma comes a drop in the cost of borrowing, all else equal. As a result, the incentives to borrow are greater. However, in equilibrium, the absence of stigma leads borrowers to quite readily file for bankruptcy and alters loan rates by enough to reduce equilibrium debt holdings.

loans conditioned only on total debt held by a household. This is a weak requirement as well, and allowing creditors to have more information would likely lead to even less bankruptcy, as credit limits would tighten in the face of negative shocks. Lastly, the decrease in stigma studied here led to counterfactual implications beyond bankruptcy rates, as debt levels fell and credit limits shrank. It is useful to note that in Athreya (2001), it was assumed that creditors could not condition on even debt holdings. Instead creditors imposed a fixed credit limit. Even in that environment, reductions in stigma could not account for observed simultaneous increases in both debt and bankruptcy. It is apparent that in this model, reductions in stigma have large contractionary effects on credit supply, and while bankruptcy rates might rise as stigma falls, the debts discharged in bankruptcy are predicted to shrink dramatically and not rise over time as they have.

On a more positive note, a promising explanation appears to be that falling transactions costs have led to greatly expanded borrowing possibilities. I found that this view has implications that are consistent with the facts on both filing rates as well as debt growth over the 1990s. The model did produce an excessive charge-off rate relative to the debt. Nonetheless, advances in intermediation, increased competition, and the democratization of credit, to the extent that they have effectively lowered borrowing costs, have led in this model to outcomes broadly consistent with the data.

REFERENCES

Aiyagari, S. Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *Quarterly Journal of Economics* 109 (3): 659–84.

Anderson, David. 1999. "A Tale of Two Stigmas." *Harvard Magazine*. (January–February). Available at http://www.harvard-magazine.com /issues/jf99/right.tale.html (accessed January 21, 2004).

Athreya, Kartik B. 2001. "The Growth of Unsecured Consumer Credit: Are We Better Off?" Federal Reserve Bank of Richmond *Economic Quarterly* 87 (3): 11–33.

______________. 2002. "Welfare Implications of the Bankruptcy Reform Act of 1999." *Journal of Monetary Economics* 49 (8): 1567–95.

______________, and Nicole B. Simpson. 2003. "Personal Bankruptcy or Public Insurance?" Federal Reserve Bank of Richmond Working Paper 03-14. Available at http://www.rich.frb.org/pubs/wpapers/ (accessed December 30, 2003).

Economic Report of the President. 2001. Available at http://w3.access.gpo.gov/usbudget/fy2002/pdf/2001_erp.pdf (accessed December 31, 2003).

Edelberg, Wendy. 2003. "Risk-Based Pricing of Interest Rates in Household Loan Markets." Federal Reserve Board Finance and Economics Discussion Series Paper 2003-62. Available at http://www.federalreserve.gov/Pubs/FEDS/2003/200362/200362pap.pdf (accessed December 31, 2003).

Ellis, Diane. 1998. "The Effect of Consumer Interest Rate Deregulation on Credit Card Volumes, Charge-Offs, and the Personal Bankruptcy Rate." *Bank Trends: Analysis of Emerging Trends in Banking* 98-05 (March), Federal Deposit Insurance Corporation. Available at http://www.fdic.gov/bank/analytical/bank/bt_9805.html (accessed December 31, 2003).

Evans, David, and Richard Schmalensee. 1999. *Paying with Plastic: The Digital Revolution in Buying and Borrowing.* Cambridge, Mass.: MIT Press.

Fay, Scott, Erik Hurst, and Michelle White. Forthcoming."The Bankruptcy Decision." *American Economic Review.*

Gross, David B., and Nicholas Souleles. 2002. "An Empirical Analysis of Personal Bankruptcy and Delinquency." *Review of Financial Studies* 15 (1): 319–47.

Livshits, Igor, James MacGee, and Michele Tertilt. 2003. "Consumer Bankruptcy: A Fresh Start." Federal Reserve Bank of Minneapolis Working Paper 617.

Moss, David A., and Gibbs A. Johnson. 1999. "The Rise of Consumer Bankruptcy: Evolution, Revolution, or Both?" *American Bankruptcy Law Journal* 73 (Spring): 311–51.

Pomykala, Joseph S. 1999. "Bankruptcy Laws: The Need for Reform." *USA Today (Magazine)* (November). Available at http://www.findarticles.com/cf_0/m1272/2654_128/57564072/p1/ article.jhtml?term=Bankruptcy (accessed January 27, 2004).

Stavins, Joanna. 2000. "Credit Card Borrowing, Delinquency, and Personal Bankruptcy." Federal Reserve Bank of Boston *New England Economic Review* (July/August): 15–30.

Sullivan, Theresa S., Elizabeth Warren, and Jay Westbrook. 2000. *The Fragile Middle Class.* New Haven: Yale University Press.

White, Michelle. 1998. "Why Don't More Households File for Bankruptcy?" *Journal of Law, Economics, and Organization* 14 (October): 205–31.

Monetary Policy and the Adjustment to Country-Specific Shocks

Margarida Duarte

The question of the optimal degree of exchange rate variability among countries has been long standing in international economics. Friedman (1953) argued in favor of flexible exchange rates: when nominal goods prices are sticky, the adjustment of the nominal exchange rate allows for the necessary relative price adjustment to a country-specific shock. Recent articles by Devereux and Engel (2003) and Corsetti and Pesenti (2001), however, show that the optimal degree of exchange rate variability between two countries subject to country-specific real shocks depends critically on the nature of price stickiness, in particular, whether prices are sticky in the currency of the producer or in the currency of the buyer.

When prices are preset in the currency of the buyer, unanticipated movements in the nominal exchange rate do not affect the price of imported goods on impact. That is, as suggested by the empirical evidence, the pass-through of exchange rate changes to consumer prices in the short run is low.[1] The findings in Devereux and Engel (2003) and Corsetti and Pesenti (2001) show that when prices are preset in the currency of the buyer (and, therefore, as suggested by the data, do not respond to movements in the exchange rate), optimal monetary policy implies that the nominal exchange rate does not respond to country-specific shocks.[2] This finding is in sharp contrast with Friedman's

I would like to thank Huberto Ennis, Andreas Hornstein, Thomas Humphrey, Diego Restuccia, and Pierre Sarte for helpful comments. Remaining errors are my own. This article does not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] Recent empirical studies have documented a low pass-through of changes in the exchange rate to consumer prices: in the short run, consumer prices respond little to changes in the nominal exchange rate. See, for example, Engel (1993, 1999) and Engel and Rogers (1996), among others.

[2] If, instead, prices are sticky in the currency of the producer, then consumer prices of imported goods change proportionally with unanticipated changes in the nominal exchange rate (com-

(1953) argument in favor of nominal exchange rate flexibility in the presence of nominal price rigidities and country-specific shocks.

In this article I study optimal monetary policy in a two-country model that features nontraded goods and in which producers in each country set prices one period in advance in the currency of the buyer. The article shows that the presence of nontraded goods has important implications for the optimal degree of nominal exchange rate variability in response to country-specific shocks.

When all goods are traded, I find that, as in Devereux and Engel (2003) and Corsetti and Pesenti (2001), both monetary authorities respond in the same manner to country-specific shocks. In the absence of nontraded goods, home and foreign agents consume the same basket of traded goods, and the nominal exchange rate does not move in response to country-specific shocks when countries follow their optimal monetary rules. As a consequence, a fixed exchange rate regime can be supported by optimal monetary policies. An important feature of the model when all goods are traded is that it implies that there are no relative price differentials across countries under a fixed exchange rate regime. There exists, however, evidence of such price differentials across countries that participate in a currency union (and, therefore, have a constant nominal exchange rate). Duarte (2003), for example, documents inflation differentials among member countries of the European Monetary Union as big as 4 percentage points.

When a set of consumption goods is nontraded and the consumption basket is distinct across countries, I show that the model is consistent with the observed relative price differentials across countries under a fixed exchange rate regime. I find, however, that in this situation a fixed exchange rate regime is not supported by optimal monetary policies since the monetary authorities choose to respond differently to country-specific shocks. That is, a flexible exchange rate regime is supported by optimal monetary policies when the model is consistent with two observations—that of relative price differentials across countries under a fixed exchange rate regime, as well as the observation of low pass-through of exchange rate changes to consumer prices.

The presence of nontraded goods has been shown to have important implications in open-economy models in dimensions other than the optimal degree of exchange rate variability. Stockman and Tesar (1995) show that nontraded goods play an important role in accounting for the properties of the international business cycle of industrialized countries. More recently, Corsetti and Dedola (2002) and Burstein, Neves, and Rebelo (2003) show that nontraded goods, used in the distribution sector, play an important role in explaining ob-

plete exchange rate pass-through). In this case, Devereux and Engel (2003) and Corsetti and Pesenti (2001) find that a flexible exchange rate regime can be supported by optimal monetary policies.

served deviations from the law of one price and low pass-through of exchange rate changes to consumer prices.

In Section 1, I present a two-country model in which agents consume traded and nontraded goods and in which prices are sticky. In the following section, I study the implications of the presence of nontraded goods for relative price differentials across countries under a fixed exchange rate regime and for the optimal response of a monetary authority seeking to maximize the expected utility of the representative agent in the country.

1. THE MODEL

In this section I develop a general equilibrium model of a world economy with two countries, denominated home and foreign, which builds upon the work of Obstfeld and Rogoff (1995). Both countries are populated by a continuum of monopolistic producers, indexed by $i \in [0, 1]$ in the home country and $i^* \in [0, 1]$ in the foreign country. Each agent produces two goods, a differentiated traded good and a differentiated nontraded good.[3] Agents consume all varieties of home and foreign-traded goods and all varieties of the local-nontraded good. In each country there is a monetary authority that prints local currency and distributes it to the individual agents through lump sum transfers.

I now describe the home economy. The foreign economy is analogous to the home economy. Foreign variables are denoted with an asterisk.

Preferences

All agents have identical preferences defined over a consumption index, real money balances, and work effort. To keep the algebra to a minimum, the lifetime expected utility of a typical home agent j is defined as

$$U_0(j) = E_0 \sum_{t=0}^{\infty} \beta^t \left(\ln c_t(j) + \chi \ln \frac{M_t(j)}{P_t} - l_t(j) \right). \tag{1}$$

The real consumption index $c_t(j)$ is defined as

$$c_t(j) = \frac{c_{T,t}(j)^{\gamma}\, c_{N,t}(j)^{1-\gamma}}{\gamma^{\gamma}(1-\gamma)^{1-\gamma}}, \tag{2}$$

where $c_{T,t}(j)$ denotes the agent's consumption index of traded goods and $c_{N,t}(j)$ denotes the agent's consumption index of nontraded goods. The consumption index of traded goods is defined as

$$c_{T,t}(j) = \frac{c_{H,t}(j)^{\eta}\, c_{F,t}(j)^{1-\eta}}{\eta^{\eta}(1-\eta)^{1-\eta}}, \tag{3}$$

[3] See Obstfeld and Rogoff (1996, 661) for a discussion of the environment where individuals, instead of firms, are the locus of monopoly power.

where $c_{H,t}(j)$ and $c_{F,t}(j)$ denote agent j's consumption index of home and foreign-traded goods, respectively.[4] Finally, the consumption indexes of home-traded goods, $c_{H,t}(j)$, foreign-traded goods, $c_{F,t}(j)$, and local-nontraded goods, $c_{N,t}(j)$, are each defined over consumption of all the varieties of each good, as

$$c_{H,t}(j) = \left[\int_0^1 c_t(h,j)^{\frac{\theta-1}{\theta}} dh\right]^{\frac{\theta}{\theta-1}}, \quad (4)$$

$$c_{F,t}(j) = \left[\int_0^1 c_t(f,j)^{\frac{\theta-1}{\theta}} df\right]^{\frac{\theta}{\theta-1}}, \quad (5)$$

and

$$c_{N,t}(j) = \left[\int_0^1 c_t(n,j)^{\frac{\theta-1}{\theta}} dn\right]^{\frac{\theta}{\theta-1}}. \quad (6)$$

In equation (4), $c_t(h,j)$ denotes agent j's consumption of home-traded variety h, $h \in [0,1]$, at date t. The terms $c_t(f,j)$ and $c_t(n,j)$ in equations (5) and (6) have analogous interpretations.

Note that in equations (2) and (3) it is assumed that the elasticity of substitution between the composite goods of home and foreign-traded varieties ($c_H(j)$ and $c_F(j)$) and the elasticity of substitution between the composite goods of traded and nontraded varieties ($c_T(j)$ and $c_N(j)$) are equal to one. In expressions (4) through (6), however, the elasticity of substitution between distinct varieties of a given good (nontraded or traded) is given by θ, which is assumed to be greater than one.[5]

Let's denote by $P_{H,t}(h)$ and $P_{F,t}(f)$ the home-currency prices of varieties h and f of the home and foreign-traded goods at date t, respectively. And let $P_{N,t}(n)$ denote the home-currency price of variety n of the local-nontraded good. The utility-based home price index, P_t, is then given by[6]

$$P_t = P_{T,t}^{\gamma} P_{N,t}^{1-\gamma}, \quad (7)$$

where the price of one unit of the composite good of all traded varieties, $P_{T,t}$, and the price of one unit of the composite good of nontraded varieties, $P_{N,t}$,

---

[4] I follow Corsetti and Pesenti (2001) and Devereux and Engel (2003) in assuming that foreign agent j^*'s consumption index of traded goods is defined as $c_{T,t}^*(j^*) = \frac{c_{H,t}^*(j^*)^{\eta} c_{F,t}^*(j^*)^{1-\eta}}{\eta^{\eta}(1-\eta)^{1-\eta}}$. That is, home and foreign agents consume the same basket of home- and foreign-traded goods. This specfication, for example, does not generate home bias.

[5] This assumption is required to ensure that an interior equilibrium with a positive level of output exists.

[6] The price index P_t is defined as the minimum expenditure required to buy one unit of the composite good c_t, given the prices of all individual varieties. The other price indexes have analogous interpretations. See the appendix for the derivation of the price indexes and the demand functions (11) through (13) presented below.

are given by

$$P_{T,t} = P_{H,t}^{\eta} P_{F,t}^{1-\eta}, \tag{8}$$

and

$$P_{N,t} = \left[\int_0^1 P_{N,t}(n)^{1-\theta} \, dn \right]^{\frac{1}{1-\theta}}. \tag{9}$$

The prices of one unit of the composite goods of home and foreign-traded varieties, in turn, are given by

$$P_{H,t} = \left[\int_0^1 P_{H,t}(h)^{1-\theta} \, dh \right]^{\frac{1}{1-\theta}}; \; P_{F,t} = \left[\int_0^1 P_{F,t}(f)^{1-\theta} \, df \right]^{\frac{1}{1-\theta}}. \tag{10}$$

For the above specification of consumption indexes, agent j's demands for variety h and f of home and foreign-traded goods are given by

$$c_t(h, j) = \eta\gamma \left(\frac{P_{H,t}(h)}{P_{H,t}} \right)^{-\theta} \frac{P_t}{P_{H,t}} c_t(j), \tag{11}$$

and

$$c_t(f, j) = (1-\eta)\gamma \left(\frac{P_{F,t}(f)}{P_{F,t}} \right)^{-\theta} \frac{P_t}{P_{F,t}} c_t(j). \tag{12}$$

The agent's demand for variety n of the nontraded good is given by

$$c_t(n, j) = (1-\gamma) \left(\frac{P_{N,t}(n)}{P_{N,t}} \right)^{-\theta} \frac{P_t}{P_{N,t}} c_t(j). \tag{13}$$

Production Technologies

The home agent j operates two technologies, one to produce a variety h of the home-traded good and the other to produce a variety n of the nontraded good. Both technologies are linear in labor. The corresponding resource constraints, which equate the quantities demanded and supplied of each variety, are

$$z_t l_{T,t}(j) \geq \int_0^1 c_t(h, i) \, di + \int_0^1 c_t^*\left(h, i^*\right) di^*, \tag{14}$$

and

$$z_t l_{N,t}(j) \geq \int_0^1 c_t(n, i) \, di, \tag{15}$$

where z_t denotes a country-specific productivity shock to both nontraded and traded technologies.[7] The term $\int_0^1 c_t(h, i) \, di$ represents aggregate demand

[7] This article concerns the adjustment to country-specific shocks, and, therefore, I abstract from sector-specific shocks.

in the home country for home variety h. The other integrals have analogous interpretations. The terms $l_{T,t}(j)$ and $l_{N,t}(j)$ denote the fraction of time that agent j allocates to production of the traded and nontraded varieties, respectively. The agent's total work effort, $l_t(j)$, is given by $l_{T,t}(j) + l_{N,t}(j)$.

Budget Constraint

Agent j holds local currency, $M_t(j)$, and trades state-contingent nominal bonds (denominated in the home currency) with foreign agents. We denote the price at date t when the state of the world is s_t of a bond paying one unit of currency at date $t+1$ if the state of the world is s_{t+1} by $Q_{s_{t+1}|s_t}$, and we denote the number of these bonds purchased by the home agent at date t by $B_{s_{t+1}}(j)$. Bond revenues received at date t when the state of the world is s_t are denoted by $B_{s_t}(j)$.

The agent's budget constraint, expressed in home-currency units, is

$$\begin{aligned} & P_t c_t(j) + \sum_{s_{t+1}} Q_{s_{t+1}|s_t} B_{s_{t+1}}(j) + M_t(j) \\ \leq \; & R_t(j) + B_{s_t}(j) + M_{t-1}(j) + T_t(j), \end{aligned} \tag{16}$$

where $P_t c_t(j)$ is nominal expenditure in consumption, $R_t(j)$ denotes sales revenues, and $T_t(j)$ denotes lump sum transfers received from the monetary authority.

Revenues from selling the traded variety h and the nontraded variety n, $R_t(j)$, are given by

$$\begin{aligned} R_t(j) \;=\; & P_{H,t}(h) \int_0^1 c_t(h,i)\,di + e_t P^*_{H,t}(h) \int_0^1 c^*_t\left(h,i^*\right) di^* \\ & + P_{N,t}(n) \int_0^1 c_t(n,i)\,di. \end{aligned}$$

In this expression, $P^*_{H,t}(h)$ denotes the foreign currency price of home-traded variety h, and $P^*_{H,t}(h) \int_0^1 c^*_t(h,i^*)\,di^*$ denotes agent j's sales revenue in the foreign country (expressed in foreign currency units). The nominal exchange rate in period t, denoted by e_t, converts foreign currency sales revenue into home-currency units.

The Agent's Problem

Agent j maximizes his expected lifetime utility (equation (1)) subject to the resource constraints for the home-traded variety h and nontraded variety n he produces (equations (14) and (15)) and his budget constraint (equation (16)), by choosing sequences of consumption, bond holdings, money holdings, and prices for the varieties h and n, taking other prices as given.

I assume that agents choose the nominal price of their traded and nontraded varieties one period in advance.[8] Moreover, I assume that producers can segment home and foreign markets and set prices for the traded variety in the currency of the buyer. Then, home producer j producing home-traded variety h and nontraded variety n chooses prices $P_{H,t}(h)$, $P^*_{H,t}(h)$, and $P_{N,t}(n)$ (where $P^*_{H,t}(h)$ is denominated in foreign currency units) at time $t-1$, taking other prices as given. The agent's problem is solved in Appendix A.

In a symmetric equilibrium, the first-order condition for consumption implies

$$\lambda_t = \frac{1}{P_t c_t}, \tag{17}$$

where λ_t, the Lagrange multiplier of the budget constraint, is the marginal utility of the (representative) agent's marginal wealth. The first-order condition for real money balances implies the money demand function

$$\frac{M_t}{P_t} = \chi c_t \frac{1+i_{t+1}}{i_{t+1}}, \tag{18}$$

where $1+i_{t+1}$ is the gross return in period $t+1$ of a riskless bond and is given by

$$\frac{1}{1+i_{t+1}} = \beta E_t \left[\frac{P_t c_t}{P_{t+1} c_{t+1}} \right]. \tag{19}$$

Finally, from the first-order conditions for state-contingent bond holdings for home and foreign agents, we obtain the risk sharing condition[9]

$$P_t c_t = e_t P^*_t c^*_t. \tag{20}$$

For the momentary utility specification in equation (1), complete risk sharing implies that nominal expenditure in consumption (when expressed in the same currency) is equalized across countries. Note that consumption of the composite good differs across the two countries only to the extent that its price (when expressed in the same currency) differs across countries, that is, when there are deviations from purchasing power parity (or $P_t \neq e_t P^*_t$).

In a symmetric equilibrium, the optimal pricing equations are

$$P_{H,t}(h) = P_{H,t} = \frac{\theta}{\theta-1} E_{t-1} \left[\frac{P_t c_t}{z_t} \right], \tag{21}$$

$$P^*_{H,t}(h) = P^*_{H,t} = \frac{\theta}{\theta-1} E_{t-1} \left[\frac{P_t c_t}{e_t z_t} \right], \tag{22}$$

---

[8] As in Corsetti and Pesenti (2001) and Devereux and Engel (2003), I abstract from a richer price adjustment setting in order to simplify the analytical solution of the model.

[9] Several recent articles have assumed complete nominal asset markets. See, for example, Chari, Kehoe, and McGrattan (2003) or Devereux and Engel (2003) for a discussion.

and

$$P_{N,t}(n) = P_{N,t} = P_{H,t}(h). \quad (23)$$

Since prices are set in advance in the currency of the buyer, it follows that, in the event of an unanticipated shock, consumer prices remain unchanged for one period. On impact, therefore, there is no pass-through of nominal exchange rate movements to consumer prices, and unanticipated changes in the nominal exchange rate cause ex-post deviations from the law of one price (that is, $P_{H,t}(h) \neq e_t P^*_{H,t}(h)$).

If, instead, agents choose prices after observing the current realization of productivity shocks, then the price rules above hold in each state of the world and not just in expectation. Note that with flexible prices, $P_{H,t}(h) = e_t P^*_{H,t}(h)$ holds every period (i.e., the law of one price holds). That is, even though firms can segment home and foreign markets, they optimally choose to charge the same price (when denominated in the same currency) in both markets when prices are flexible.

Monetary Authority

The monetary authority prints money and rebates the seigniorage revenue to agents through lump sum transfers. Its budget constraint is

$$\int_0^1 (M_t(j) - M_{t-1}(j))\, dj = \int_0^1 T_t(j)\, dj.$$

I assume that the monetary authority controls the nominal interest rate and supplies the amount of nominal money balances demanded. I follow Corsetti and Pesenti (2001) in characterizing monetary policy in each country by the reciprocal of the marginal utility of the representative agent's nominal wealth, $\mu \equiv \frac{1}{\lambda}$. In equilibrium, the marginal utility of wealth is given by equation (17), and the nominal interest rate (equation (19)) can be expressed as

$$\frac{1}{1 + i_{t+1}} = \beta E_t \left[\frac{\mu_t}{\mu_{t+1}} \right].$$

Given a time path for μ, there is a corresponding sequence of home nominal interest rates. Note that, for an unchanged $E_t \left[\frac{1}{\mu_{t+1}} \right]$, an expansionary monetary policy shock (higher μ_t and higher nominal expenditure $P_t c_t$ in equilibrium) is associated with a lower nominal interest rate i_{t+1} and, therefore (from equation (18)), with higher money balances demanded.

The Solution of the Model

The solution of the model can be easily obtained in closed form by expressing all endogenous variables as functions of real shocks (z_t and z^*_t) and monetary stances (μ_t and μ^*_t). The solution of the model is derived in Appendix B.

From equation (20), it follows that the nominal exchange rate is given by

$$e_t = \frac{\mu_t}{\mu_t^*}. \tag{24}$$

Total consumption in the home country is given by

$$c_t = \frac{\mu_t}{P_t}, \tag{25}$$

where the price index P_t is given by

$$P_t = \frac{\theta}{\theta - 1} E_{t-1}\left(\frac{\mu_t}{z_t}\right)^{\eta\gamma+(1-\gamma)} E_{t-1}\left(\frac{\mu_t}{z_t^*}\right)^{(1-\eta)\gamma}. \tag{26}$$

Total labor effort in the home country is given by

$$l_t = \frac{1}{z_t \frac{\theta}{\theta-1}} \left[\frac{(\gamma\eta + (1-\gamma))\,\mu_t}{E_{t-1}\left(\frac{\mu_t}{z_t}\right)} + \frac{\gamma\eta\mu_t^*}{E_{t-1}\left(\frac{\mu_t^*}{z_t}\right)} \right]. \tag{27}$$

Note that consumption in the home country is independent of (contemporaneous) changes in the nominal exchange rate when prices are preset in the buyer's currency. Therefore, consumption in the home country is not affected by foreign monetary policy, μ_t^*. Note also that, for given μ_t and μ_t^*, real shocks do not have a contemporaneous impact on consumption (and, therefore, output), only affecting labor effort in the country where the shock occurs. In response to a positive productivity shock in the home country, home agents produce the same quantity of traded and nontraded goods with less hours of work.

It is also useful to characterize total consumption and labor allocations when prices are flexible. In this case, total consumption and total labor effort in the home country (denoted with the superscript fl) are given by

$$c_t^{fl} = \frac{z_t^{\gamma\eta+(1-\gamma)} z_t^{*(1-\eta)\gamma}}{\frac{\theta}{1-\theta}} \tag{28}$$

and

$$l^{fl} = \frac{2\gamma\eta + (1-\gamma)}{\frac{\theta}{1-\theta}}. \tag{29}$$

With flexible prices, total consumption depends only on real shocks and is independent of monetary policy. In response to a positive productivity shock in the home country, total consumption increases more in the home country than in the foreign country. Since this shock affects total consumption differently in the two countries, it is associated with an equilibrium real interest rate differential across the two countries. Note also that, since foreign-traded goods become relatively more expensive than home goods (traded and nontraded), agents substitute consumption toward goods produced in the home country and away from goods produced in the foreign country.[10] Labor effort in each sector remains unchanged.

2. MONETARY POLICY

In this section, I start by studying the implications of nontraded goods for the nature of relative price differentials across countries under a fixed exchange rate regime. I then turn to the implications of the presence of nontraded goods for the optimal response of monetary policy to country-specific shocks.

Relative Price Differentials

Under a fixed nominal exchange rate regime, it follows from equation (24) that home and foreign monetary stances, μ and μ^*, are proportional. That is, $\mu_t = \overline{e}\mu_t^*$, where $\overline{e}$ is the fixed level of the nominal exchange rate.[11] The price level in the home country, P_t, is given by equation (26) while the price level in the foreign country (expressed in foreign currency units), P_t^*, is given by

$$P_t^* = \frac{1}{\overline{e}}\frac{\theta}{\theta-1}E_{t-1}\left[\frac{\mu_t}{z_t}\right]^{\eta\gamma} E_{t-1}\left[\frac{\mu_t}{z_t^*}\right]^{(1-\eta)\gamma+1-\gamma}.$$

The relative price across countries is then given by

$$\frac{\overline{e}P_t^*}{P_t} = \left(\frac{E_{t-1}\left[\frac{\mu_t}{z_t^*}\right]}{E_{t-1}\left[\frac{\mu_t}{z_t}\right]}\right)^{1-\gamma}.$$

Note that when $\gamma \rightarrow 1$ (that is, when agents do not consume local-nontraded goods and consume the same basket of traded goods), the relative price across countries is constant. This feature results from the fact that home and foreign agents consume the exact same basket of goods, and the nominal exchange rate is constant. Therefore, without nontraded goods, the model cannot account for the observed relative price differentials across countries when the nominal exchange rate is fixed.[12]

In the presence of nontraded goods ($\gamma < 1$), home and foreign agents consume distinct baskets of goods, and country-specific shocks lead to relative price differentials across countries (one period after the shock) under a fixed

[10] In the home country, for example, it follows from the pricing rules and demand functions presented above that the relative price of home-traded goods in terms of foreign-traded goods is $\frac{P_{H,t}}{P_{F,t}} = \frac{z_t^*}{z_t}$, and the ratio of home to foreign-traded goods consumed is $\frac{c_{H,t}}{c_{F,t}} = \frac{\eta}{1-\eta}\frac{z_t}{z_t^*}$.

[11] Note that a fixed exchange rate regime can be interpreted as a monetary policy prescription where home and foreign monetary policy stances are proportional. Later in this section, I will show under which conditions such a prescription is optimal.

[12] See, for example, Duarte (2003) for empirical evidence on relative price differentials across countries in the European Union.

exchange rate regime. With nontraded goods, the model is consistent with observed relative price across countries when the exchange rate is fixed.

Country-Specific Shocks

I now turn to the implications of the presence of nontraded goods for the optimal response of monetary policy to country-specific shocks. I follow Corsetti and Pesenti (2001) and Devereux and Engel (2003) and assume that the monetary authority in each country commits to preannounced state-contingent monetary stances, $\{\mu(s_\tau), \mu^*(s_\tau)\}_{\tau=t}^{\infty}$, chosen to maximize the (non-monetary) expected utility of the country's representative agent and taking the monetary policy rule of the other country as given.[13] That is, the monetary authority in the home country solves

$$\max_{\{\mu(s_{t+\tau})\}_{\tau=0}^{\infty}} E_{t-1}\left[\sum_{\tau=t}^{\infty}\beta^{\tau-t}\left(\ln c(s_\tau) - l(s_\tau)\right)\right], \tag{30}$$

taking $\{\mu^*(s_{t+\tau}), z(s_{t+\tau}), z^*(s_{t+\tau})\}_{\tau=0}^{\infty}$ as given. It is shown in the Appendix C that the optimal monetary stances of the home and foreign monetary authorities are given by

$$\mu_t = \left[\frac{\eta\gamma + (1-\gamma)}{z_t E_{t-1}\left(\frac{\mu_t}{z_t}\right)} + \frac{(1-\eta)\gamma}{z_t^* E_{t-1}\left(\frac{\mu_t}{z_t^*}\right)}\right]^{-1}, \tag{31}$$

and

$$\mu_t^* = \left[\frac{\eta\gamma}{z_t E_{t-1}\left(\frac{\mu_t^*}{z_t}\right)} + \frac{(1-\eta)\gamma + (1-\gamma)}{z_t^* E_{t-1}\left(\frac{\mu_t^*}{z_t^*}\right)}\right]^{-1}. \tag{32}$$

Note first that, in the absence of nontraded goods (that is, $\gamma \to 1$), the two monetary authorities choose to respond equally to country-specific shocks. Therefore, when home and foreign agents consume exactly the same basket of goods, the nominal exchange rate does not respond to country-specific productivity shocks: a fixed nominal exchange rate regime is consistent with the optimal monetary policy rules. This result replicates the findings in Devereux and Engel (2003) and Corsetti and Pesenti (2001). Note, however, that as we have seen, in this case the model misses on an important aspect of the empirical evidence by not implying relative price differentials across countries.

When agents consume local-nontraded goods (that is, $\gamma < 1$), consumption baskets differ across countries, and the rules above imply that the home

[13] As it is standard in the literature, I assume that governments ignore the utility from real balances and consider optimal policies when $\chi \to 0$.

and foreign monetary authorities choose to respond differently to country-specific productivity shocks. Therefore, the nominal exchange rate responds to country-specific productivity shocks (equation (24)) and a fixed nominal exchange rate regime is not consistent with the optimal monetary rules. Furthermore, consistent with the evidence, if the two countries adopt a fixed exchange regime, the model implies relative price differentials across countries.

In response to a positive productivity shock in the home country (and starting from a symmetric equilibrium), rules (31) and (32) require a larger expansionary monetary policy (higher μ) in the home country than in the foreign country when $\gamma < 1$. These responses are associated with a depreciation of the nominal exchange rate (equation (24)). As in the case with flexible prices, total consumption increases more in the home country than in the foreign country in response to a positive real shock in the home country. The terms of trade, however, are not affected by this shock (as they would be if prices were flexible) since prices are preset in the buyer's currency. That is, there is no consumption substitution toward goods produced in the home country: consumption of all goods in a given country increases in the same proportion.

In a fixed exchange rate regime, identical responses by both home and foreign monetary authorities cannot generate the distinct consumption paths across countries associated with a country-specific shock. This result follows from the fact that countries share a common nominal interest rate and prices are preset. Therefore, the optimal responses by the monetary authorities, which generate the distinct response of consumption across countries, require independent monetary policies and, hence, an adjustable nominal exchange rate. This result is consistent with Friedman's (1953) case in favor of nominal exchange rate flexibility in the presence of nominal price rigidities and country-specific shocks.

3. CONCLUSION

In this article I develop a two-country general equilibrium model with traded and nontraded goods where goods prices are set one period in advance in the currency of the buyer. The monetary authority in each country follows a state-contingent monetary policy rule that maximizes the expected utility of the representative agent.

I show that the presence of nontraded goods has important implications for the nature of price differentials across countries under a fixed exchange rate regime and for the optimal degree of nominal exchange rate variability in response to country-specific shocks. When there are nontraded goods, agents in different countries consume different baskets of goods and the optimal monetary policy implies that the nominal exchange rate varies in response to country-specific shocks. In contrast, when all goods are traded, agents in different countries consume the same basket of goods, and the optimal monetary policy implies that the nominal exchange rate is constant in response to country-specific shocks.

The results in this article indicate the importance of observed price differentials across countries in the evaluation of alternative exchange rate regimes. The results indicate that the existence of nontraded goods imposes a welfare cost to countries in a currency area that face country-specific shocks.

APPENDIX A: THE AGENT'S PROBLEM

Intratemporal problem

Given the consumption index (4), the utility-based price index P_H is the price of c_H that solves

$$\min_{c(h,j)} \int_0^1 c(h, j)\, P_H(h)\, dh$$

s.t.

$$c_H(j) = \left[\int_0^1 c(h, j)^{\frac{\theta-1}{\theta}}\, dh\right]^{\frac{\theta}{\theta-1}} = 1.$$

The equation for P_H in (10) in the text is the solution to this problem. The other price indexes are obtained from analogous problems.

To solve for the demand for individual variety h, consider the problem of allocating a given level of nominal expenditure X_H among varieties of home-traded good:

$$\max_{c(h,j)} \left[\int_0^1 c(h, j)^{\frac{\theta-1}{\theta}}\, dh\right]^{\frac{\theta}{\theta-1}}$$

s.t.

$$\int_0^1 c(h, j)\, P_H(h)\, dh = X_H.$$

From the first-order conditions for any pair of varieties h and h' we have

$$\begin{aligned}
\frac{c(h,j)}{c(h',j)} &= \left(\frac{P_H(h)}{P_H(h')}\right)^{-\theta} \Leftrightarrow \\
c(h,j)^{\frac{\theta-1}{\theta}} P_H\left(h'\right)^{1-\theta} &= c\left(h',j\right)^{\frac{\theta-1}{\theta}} P_H(h)^{1-\theta} \Rightarrow \\
\left(\int_0^1 c(h,j)^{\frac{\theta-1}{\theta}} P_H\left(h'\right)^{1-\theta} dh'\right)^{\frac{\theta}{\theta-1}} &= \\
\left(\int_0^1 c\left(h',j\right)^{\frac{\theta-1}{\theta}} P_H(h)^{1-\theta} dh'\right)^{\frac{\theta}{\theta-1}} &\Leftrightarrow \\
c(h,j)\underbrace{\left(\int_0^1 P_H\left(h'\right)^{1-\theta} dh'\right)^{\frac{\theta}{\theta-1}}}_{P_H^{-\theta}} &= \\
P_H(h)^{-\theta}\underbrace{\left(\int_0^1 c\left(h',j\right)^{\frac{\theta-1}{\theta}} dh'\right)^{\frac{\theta}{\theta-1}}}_{c_H(j)} &\Leftrightarrow \\
c(h,j) &= \left(\frac{P_H(h)}{P_H}\right)^{-\theta} c_H(j). \qquad (33)
\end{aligned}$$

Note that by rearranging equation (33) we obtain $\int_0^1 c(h,j)\, P_H(h)\, dh = P_H c_H(j)$.

Following analogous derivations, we obtain

$$c_H(j) = \eta \frac{P_T}{P_H} c_T(j)$$

and

$$c_T(j) = \gamma \frac{P}{P_T} c(j).$$

Combining these two expressions with equation (33) yields equation (11) in the text. Equations (12) and (13) are obtained in a similar way.

Intertemporal problem

The problem of home agent j, who produces traded variety h and nontraded variety n, is

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \left(\ln c_t(j) + \chi \ln \frac{M_t(j)}{P_t} - l_t(j)\right),$$

subject to the budget constraint

$$\begin{aligned} & P_t c_t(j) + \sum_{s_{t+1}} Q_{s_{t+1}|s_t} B_{s_{t+1}}(j) + M_t(j) \\ \leq \; & R_t(j) + B_{s_t}(j) + M_{t-1}(j) + T_t(j), \end{aligned}$$

where revenues, $R_t(j)$, and labor effort, $l_t(j)$, are given by

$$\begin{aligned} R_t(j) &= P_{H,t}(h) \int_0^1 c_t(h,i)\,di + e_t P_{H,t}^*(h) \int_0^1 c_t^*\left(h,i^*\right) di^* \\ &\quad + P_{N,t}(n) \int_0^1 c_t(n,i)\,di, \end{aligned}$$

and

$$\begin{aligned} l_t(j) &= l_{T,t}(j) + l_{N,t}(j) \\ &= \frac{1}{z_t}\left(\int_0^1 c_t(h,i)\,di + \int_0^1 c_t^*\left(h,i^*\right) di^* + \int_0^1 c_t(n,i)\,di\right). \end{aligned}$$

The first-order conditions with respect to $c_t(j)$, $M_t(j)$, and $B_{s_{t+1}}(j)$ are, respectively,

$$\frac{1}{c_t(j)} = \lambda_t(j) P_t, \tag{34}$$

$$\frac{\chi}{M_t(j)} = \lambda_t(j) - \beta E_t\left[\lambda_{t+1}(j)\right], \tag{35}$$

and

$$\lambda_{s_t}(j)\, Q_{s_{t+1}|s_t} = \beta\pi\left(s_{t+1}|s_t\right) \lambda_{s_{t+1}}(j), \tag{36}$$

where $\pi(s_{t+1}|s_t)$ is the conditional probability of event s_{t+1}, given s_t.

Recall that pricing decisions are made before the realization of period t shocks. Therefore, the first-order conditions with respect to $P_{H,t}(h)$, $P_{H,t}^*(h)$, and $P_{N,t}(n)$ are, respectively,

$$E_{t-1}\left[-\underbrace{(-\theta)\frac{\int_0^1 c_t(h,i)\,di}{z_t P_{H,t}(h)}}_{\frac{\partial l_t(j)}{\partial P_{H,t}(h)}} + \lambda_t(j) \underbrace{(1-\theta)\int_0^1 c_t(h,i)\,di}_{\frac{\partial R_t(j)}{\partial P_{H,t}(h)}} \right] = 0, \tag{37}$$

$$E_{t-1}\left[-\theta\frac{\int_0^1 c_t^*(h,i)\,di}{z_t P_{H,t}^*(h)} + \lambda_t(j)(1-\theta)\int_0^1 c_t^*(h,i)\,di \right] = 0, \tag{38}$$

and

$$E_{t-1}\left[-\theta\frac{\int_0^1 c_t(n,i)\,di}{z_t P_{N,t}(n)} + \lambda_t(j)(1-\theta)\int_0^1 c_t(n,i)\,di \right] = 0. \tag{39}$$

All agents within one country solve identical problems and therefore make identical choices (even though they produce differentiated varieties of the traded and nontraded goods). In a symmetric equilibrium in which all individual variables are identical, it follows that aggregate quantities are equal to per capita quantities (since the measure of agents is one in both countries) and that the price indexes $P_{H,t}$, $P^*_{H,t}$, and $P_{N,t}$ equal the price of its varieties ($P_{H,t}(h)$, $P^*_{H,t}(h)$, and $P_{N,t}(n)$, respectively). That is, per capita consumption of variety h in the home country is $c_t(h,i), \forall i$, which is equal to aggregate consumption of this variety: $c_{H,t} \equiv \int_0^1 c_t(h,i)\,di = c_t(h,i)$. Per capita total consumption is $c_t(i)$, which is equal to aggregate total consumption: $c_t \equiv \int_0^1 c_t(i)\,di = c_t(i)$. In what follows, I focus on the symmetric equilibrium and therefore drop the index for the agent.

Combining equations (36) and (34) implies that

$$Q_{s_{t+1}|s_t} = \beta\pi(s_{t+1}|s_t)\frac{P_{s_t}c_{s_t}}{P_{s_{t+1}}c_{s_{t+1}}}. \tag{40}$$

Let's denote the gross return in period $t+1$ of a riskless bond as $1+i_{t+1}$. Note that the gross return $1+i_{t+1}$ is equal to the reciprocal of the price in period t of a bond paying one unit of home currency in period $t+1$ with certainty, Q_{t+1}. Since asset markets are complete, it follows that $Q_{t+1} = E_t\left[Q_{s_{t+1}|s_t}\right]$. And from (40), it follows that

$$\frac{1}{1+i_{t+1}} = Q_{t+1} = \beta E_t\left[\frac{P_t c_t}{P_{t+1}c_{t+1}}\right]. \tag{41}$$

The first-order condition for money, equation (35), can be written as

$$\chi\frac{P_t c_t}{M_t} = 1 - \beta E_t\left[\frac{P_t c_t}{P_{t+1}c_{t+1}}\right],$$

in a symmetric environment. Combining this expression with equation (41) yields the money demand equation (18) in the text.

I now turn to the pricing equations (37) through (39). Note from equation (11) that, in a symmetric equilibrium, expenditure in the composite good of home-traded varieties is a constant share (given by $\eta\gamma$) of total expenditure, that is, $P_{H,t}c_{H,t} = \eta\gamma P_t c_t$. Equation (37) can be simplified by making use of equation (34) and this relationship between total expenditure and expenditure in the composite good of home-traded varieties. Taking into account that $P_{H,t}(h)$ is known as off $t-1$, equation (37) can be rewritten as

$$\frac{\theta}{P_{H,t}(h)^2}E_{t-1}\left[\frac{\eta\gamma P_t c_t}{z_t}\right] = \frac{(\theta-1)}{P_{H,t}(h)}\eta\gamma,$$

from where equation (21) in the text directly follows. Equations (22) and (23) can be obtained in a similar fashion.

The foreign agent solves a similar problem to the one of the home agent. Note, however, that since bonds are denominated in home currency, the budget constraint of foreign agent j^* (expressed in foreign currency units) is

$$\begin{aligned} & P_t^* c_t^* \left(j^*\right) + \sum_{s_{t+1}} \frac{Q_{s_{t+1}|s_t}}{e_t} B_{s_{t+1}}^* \left(j^*\right) + M_t^* \left(j^*\right) \\ \leq \quad & R_t^* \left(j^*\right) + \frac{B_{s_t}^* \left(j^*\right)}{e_t} + M_{t-1}^* \left(j^*\right) + T_t^* \left(j^*\right). \end{aligned}$$

The first-order condition with respect to bond holdings is (in a symmetric equilibrium)

$$Q_{s_{t+1}|s_t} = \beta \pi \left(s_{t+1}|s_t\right) \frac{e_{s_t} P_{s_t}^* c_{s_t}^*}{e_{s_{t+1}} P_{s_{t+1}}^* c_{s_{t+1}}^*}.$$

Combining this equation with equation (40) implies that $\frac{P_{s_{t+1}} c_{s_{t+1}}}{e_{s_{t+1}} P_{s_{t+1}}^* c_{s_{t+1}}^*} = \frac{P_{s_t} c_{s_t}}{e_{s_t} P_{s_t}^* c_{s_t}^*}$. By iterating this equation backwards we obtain $\frac{P_{s_t} c_{s_t}}{e_{s_t} P_{s_t}^* c_{s_t}^*} = \frac{P_0 c_0}{e_0 P_0^* c_0^*}$. Assuming an equal wealth distribution across countries at date 0 implies $\frac{P_0 c_0}{e_0 P_0^* c_0^*} = 1$, which gives equation (24) in the text.

APPENDIX B: SOLUTION OF THE MODEL

To write real aggregate consumption as a function of real shocks (z_t and z_t^*) and monetary stances (μ_t and μ_t^*), note that, in equilibrium, $P_t c_t = \mu_t$. Using equations (7) and (8), and the pricing equations (21) through (23), we can write the price level P_t as

$$P_t = \Bigg(\underbrace{\left(\frac{\theta}{\theta-1} E_{t-1}\left(\frac{\mu_t}{z_t}\right)\right)^{\eta} \left(\frac{\theta}{\theta-1} E_{t-1}\left(\frac{\mu_t}{z_t^*}\right)\right)^{1-\eta}}_{P_{T,t}} \Bigg)^{\gamma}$$

$$\Bigg(\underbrace{\frac{\theta}{\theta-1} E_{t-1}\left(\frac{\mu_t}{z_t}\right)}_{P_{N,t}} \Bigg)^{1-\gamma} = \frac{\theta}{\theta-1} \left(E_{t-1}\left(\frac{\mu_t}{z_t}\right)\right)^{\eta\gamma+(1-\gamma)}$$

$$\left(E_{t-1}\left(\frac{\mu_t}{z_t^*}\right)\right)^{(1-\eta)\gamma}.$$

Total consumption is then given by

$$\begin{aligned} c_t &= \frac{\mu_t}{P_t} \\ &= \frac{\mu_t}{\frac{\theta}{\theta-1}\left(E_{t-1}\left(\frac{\mu_t}{z_t}\right)\right)^{\eta\gamma+(1-\gamma)}\left(E_{t-1}\left(\frac{\mu_t}{z_t^*}\right)\right)^{(1-\eta)\gamma}}. \end{aligned}$$

The expression for foreign aggregate consumption can be obtained in a similar way.

From the market clearing conditions for home-traded goods and nontraded goods, it follows that labor effort in the home-traded and nontraded sectors can be written as, respectively,

$$\begin{aligned} l_{T,t} &= \frac{c_{H,t} + c^*_{H,t}}{z_t}, \\ &= \frac{1}{z_t}\frac{\gamma\eta}{\frac{\theta}{\theta-1}}\left(\frac{\mu_t}{E_{t-1}\left(\frac{\mu_t}{z_t}\right)} + \frac{\mu_t^*}{E_{t-1}\left(\frac{\mu_t^*}{z_t}\right)}\right), \end{aligned}$$

and

$$\begin{aligned} l_{N,t} &= \frac{c_{N,t}}{z_t}, \\ &= \frac{1}{z_t}\frac{1-\gamma}{\frac{\theta}{\theta-1}}\frac{\mu_t}{E_{t-1}\left(\frac{\mu_t}{z_t}\right)}, \end{aligned}$$

where the second equalities follow from substituting for the demand functions. Total labor effort is simply given by $l_t = l_{T,t} + l_{N,t}$.

If, instead of setting prices before the realization of uncertainty, producers set prices after observing the current realization of productivity and monetary stance shocks (flexible prices) then, as noted in the text, the pricing equations (21) through (23) hold in every state of the world and not only in expectation. That is, with flexible prices, the prices of nontraded goods and home-traded goods at home and abroad are given by

$$P_{N,t} = P_{H,t} = \frac{\theta}{\theta-1}\frac{\mu_t}{z_t},$$

and

$$P^*_{H,t} = \frac{P_{H,t}}{e_t}.$$

The above expressions for aggregate consumption and total labor effort in the home country simplify to equations (28) and (29) in the text.

APPENDIX C: OPTIMAL POLICIES

The monetary authority in each country commits to state-contingent monetary stances $\mu(s_t)$ and $\mu^*(s_t)$, chosen to maximize the (non-monetary) expected utility of the country's representative agent, taking the monetary policy rule of the other country as given. The problem of the home monetary authority is

$$\max_{\{\mu(s_{t+\tau})\}_{\tau=0}^{\infty}} E_{t-1}\left[\sum_{\tau=t}^{\infty}\beta^{\tau-t}\left(\ln c(s_\tau)-l(s_\tau)\right)\right], \tag{42}$$

taking $\{\mu^*(s_{t+\tau}), z(s_{t+\tau}), z^*(s_{t+\tau})\}_{\tau=0}^{\infty}$ as given.

Let's focus on the choice of $\mu(s_t)$ and rewrite (42) as

$$\max_{\{\mu(s_{t+\tau})\}_{\tau=0}^{\infty}} \sum_{s_t}\pi(s_t|s_{t-1})\left[\ln c(s_t)-l(s_t)\right]+E_{t-1}$$
$$\left[\sum_{\tau=t+1}^{\infty}\beta^{\tau-t}\left(\ln c(s_\tau)-l(s_\tau)\right)\right].$$

First, note that $E_{t-1}[l(s_t)]=l^{fl}(s_t)$. Second, note that, by using equation (25), we have

$$\begin{aligned}\ln c_{s_t} &= \ln\mu_{s_t}-(\gamma\eta+1-\gamma)\ln\left(\sum_{s_t}\pi(s_t|s_{t-1})\frac{\mu(s_t)}{z(s_t)}\right)\\ &\quad -\gamma(1-\eta)\ln\left(\sum_{s_t}\pi(s_t|s_{t-1})\frac{\mu(s_t)}{z^*(s_t)}\right)+a,\end{aligned}$$

where a is a constant.

The first-order condition of the monetary authority's problem with respect to $\mu(\widetilde{s}_t)$ is

$$\frac{\pi(\widetilde{s}_t|s_{t-1})}{\mu(\widetilde{s}_t)}-\sum_{s_t}\pi(s_t|s_{t-1})$$
$$\left[\frac{(\gamma\eta+1-\gamma)\frac{\pi(\widetilde{s}_t|s_{t-1})}{z(\widetilde{s}_t)}}{\sum_{s_t}\pi(s_t|s_{t-1})\frac{\mu(s_t)}{z(s_t)}}+\frac{\gamma(1-\eta)\frac{\pi(\widetilde{s}_t|s_{t-1})}{z^*(\widetilde{s}_t)}}{\sum_{s_t}\pi(s_t|s_{t-1})\frac{\mu(s_t)}{z^*(s_t)}}\right]=0.$$

Since the term in square brackets is independent of s_t and $\sum_{s_t}\pi(s_t|s_{t-1})=1$, we can rewrite the above first-order condition as

$$\frac{1}{\mu(\widetilde{s}_t)}=\frac{\gamma\eta+1-\gamma}{z(\widetilde{s}_t)\sum_{s_t}\pi(s_t|s_{t-1})\frac{\mu(s_t)}{z(s_t)}}+\frac{\gamma(1-\eta)}{z^*(\widetilde{s}_t)\sum_{s_t}\pi(s_t|s_{t-1})\frac{\mu(s_t)}{z^*(s_t)}},$$

which yields equation (31) in the text.

REFERENCES

Burstein, Ariel, João Neves, and Sérgio Rebelo. 2003. "Distribution Costs and Real Exchange Rate Dynamics During Exchange-Rate Based Stabilizations." *Journal of Monetary Economics* 50: 1189–214.

Corsetti, Giancarlo, and Luca Dedola. 2002. "Macroeconomics of International Price Discrimination." Manuscript, European University Institute and European Central Bank.

Corsetti, Giancarlo, and Paolo Pesenti. 2001. "International Dimensions of Optimal Monetary Policy." National Bureau of Economic Research Working Paper 8230.

Devereux, Michael, and Charles Engel. 2003. "Monetary Policy in the Open Economy Revisited: Price Setting and Exchange Rate Flexibility." *The Review of Economic Studies* 70 (October): 765.

Duarte, Margarida. 2003. "The Euro and Inflation Divergence in Europe." Federal Reserve Bank of Richmond *Economic Quarterly* 89 (Summer): 53–70.

Engel, Charles. 1993. "Real Exchange Rates and Relative Prices: An Empirical Investigation." *Journal of Monetary Economics* 32: 35–50.

__________. 1999. "Accounting for U.S. Real Exchange Rate Changes." *Journal of Political Economy* 107: 507–38.

__________, and John Rogers. 1996. "How Wide is the Border?" *American Economic Review* 86: 1112–25.

Friedman, Milton. 1953. "The Case for Flexible Exchange Rates." in *Essays in Positive Economics*. Chicago: The University of Chicago Press: 157-203.

Obstfeld, Maurice, and Kenneth Rogoff. 1995. "Exchange Rate Redux." *Journal of Political Economy* 103 (June): 624–60.

__________. 1996. *Foundations of International Economics*. Cambridge: The MIT Press.

Stockman, Alan, and Linda Tesar. 1995. "Tastes and Technology in a Two-Country Model of the Business Cycle: Explaining International Comovements." *American Economic Review* 85 (March): 168–85.

Some Recent Trends in Commercial Banking

Huberto M. Ennis

In this article I review some recent trends in the evolution of U.S. commercial banks. The banking industry has experienced a series of significant transformations in the last two or three decades. Among the most important of them is the change in the type of organizations that dominate the landscape. Since the eighties, banks have increased the scope and scale of their activities, and several banks have become very large institutions with a presence in multiple regions of the country. After this long period of transformations, now is a good time to stop and look back at the changes that have occurred.

Reviewing these trends over the last thirty years may also help to put in perspective the behavior of the banking system during the 2001 recession. We have not had many recessions since the major transformations in banking happened. In fact, the only other recession took place in 1991–1992 and found the banking system in the middle of the resolution of a widespread crisis. After two years of slow recovery, this is an appropriate time to assess how the new banking system behaved during the last recession. For such evaluation, it is useful to have a long-run (thirty year) perspective on the direction of change in the relevant variables. This perspective is the focus of the present article.

This study concentrates only on commercial banks, most of which are part of a larger company, a bank holding company. Sometimes, more than one commercial bank belongs to the same bank holding company. These banks are called sister banks. In general, sister banks tend to be managed as different branches of a single bank rather than as independent banks. In fact,

Research Department, Federal Reserve Bank of Richmond, huberto.ennis@rich.frb.org. I would like to thank Margarida Duarte, Tom Humphrey, Jennifer Sparger, John Walter, John Weinberg, and Steven Zunic for their useful comments. I would also like to thank Fan Ding, Dan Herlihy, and Sam Malek for providing able research assistance. All errors are, of course, my own. The views expressed here do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

until the mid-1980s, the creation of sister banks was partly a response to strict regulations on branching. In this sense, looking at sister banks as different banks can be somewhat misleading, even from the perspective of regulators, since sister banks are subject to cross-guarantee provisions by which one bank can be liable for its sister bank's losses. However, only 30 percent of U.S. banks belong to multi-bank holding companies. Furthermore, the parent organization is subject to limited liability protection rules with respect to the losses in the bank. These rules make the financial situation of the individual banks still important for regulatory purposes, to the extent that the losses associated with a bank's failure may still be transferred to the insurance fund, even when the bank holding company does not go bankrupt. For these reasons, and since the data on commercial banks is readily available, it seems appropriate to focus on the population of commercial banks to identify the relevant major trends in the industry.

When possible, I will discuss the evolution of the industry from 1970 to the present. Boyd and Gertler (1993) reviewed the trends in U.S. commercial banking since the fifties. My analysis complements theirs, as I extend the data period to include the last ten years (see also Carlson and Perli 2002). In some cases, data availability allows me to go back only to 1984 when the procedures used by banks to report information to the regulators (The Call Reports) were thoroughly revised.

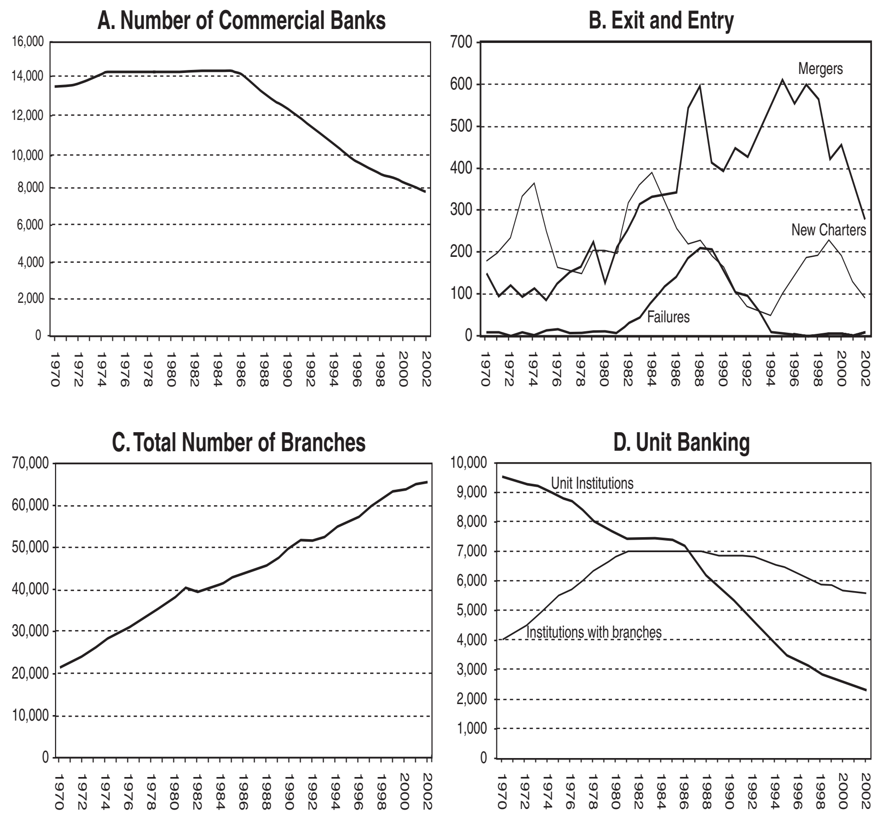
The article is organized as follows. In Section 1, I provide a quick overview of the changes in the structure of the industry. In Section 2, I study the major trends in the balance sheet of all commercial banks, taken as a group. When informative, I divide the banks into four different groups, according to their size measured by total assets. In Section 3, I perform a similar analysis of the evolution of the income statement of banks. Finally, I reserve Section 4 for some conclusions.

1. THE INDUSTRY STRUCTURE

For the period under study, one of the main changes in the structure of the banking industry has been a movement toward consolidation. An entire literature has developed around this subject (see Ennis [2001] for a summary discussion and further references). Here I present only some major trends that reflect this process.

In general, advances in communication have made the U.S. economy much more integrated at the national level. The tendency to run national operations (as opposed to regional) is quite evident in the retail sector, for example. The consequences of this trend to banking come directly and indirectly. The direct effect originates in the fact that a large portion of the activities of banks, such as the provision of checking accounts and other payment instruments, is retail in nature and hence subject to the same forces to become national. More

Figure 1 The Number of Commercial Banks



Source: Federal Deposit Insurance Corporation (Annual Data for FDIC-Insured Commercial Banks: Tables CB01, CB02, CB03).

indirectly, it is also the case that some of the major customers of banks are retailers, which, having acquired national presence, are now more likely to benefit from a relationship with a nationwide bank. The natural response has then been for banks to move toward consolidation.

At the beginning of the eighties, the number of commercial banks was more than 14,000, but by 2002, this number was less than 8,000 (see Panel A of Figure 1). It is striking to see the sharp change in 1985 to a significant downward trend in the number of commercial banks. Several factors are associated with this change. Panel B of Figure 1 shows that the number of mergers increased during the eighties and nineties. It should be noticed, however, that this time series is quite volatile. The number of new charters also decreased during the eighties, but the creation of new banks went back up to historical levels during the second half of the nineties. Finally, the number of bank failures increased significantly during the second half of the eighties but again went back to normal levels (or even lower) after 1994.[1]

The late eighties and early nineties were a period of crisis in the commercial banking industry, clearly reflected by the increase in the number of bank failures. However, the time series for the number of commercial banks does not seem to reflect this abnormal period: the number of banks is decreasing at an almost constant rate since 1985 (see Panel A of Figure 1). One possible interpretation for this lack of response in the number of banks during the crisis is that the trend in this variable is dictated by changes in long-run factors like technological progress (see Broaddus [1998] for a similar argument). In this sense, we can think that at each point in time there is a target number of banks in the economy which is determined by the technology of production, the factors influencing demand, and, of course, government regulations. Under this interpretation, the target number of banks has been subject to a relatively constant decreasing trend that started in the mid-eighties, and the crisis that occurred during the late eighties and early nineties was an abrupt change affecting only who were the participants in the industry (but not how many participants the industry should have). In other words, during the crisis some banks failed and some new banks were created, but this process had no significant effect on the underlying speed of consolidation.

In spite of the decreasing trend in the number of banks, it is interesting to note that the number of branches has actually been increasing (and at about the historic trend) during this consolidation period (see Panel C in Figure 1). One of the fears commonly associated with a move toward consolidation is that competition may be reduced as fewer and bigger banks dominate the market. However, the increase in the number of branches may be an indication that the level of competition in regional markets has not significantly decreased.[2]

Panel D in Figure 1 shows the declining trend in the number of unit banks. In the seventies, unit banks were being replaced by institutions with branches. After the mid-eighties, both the number of unit banks and the number of institutions with branches trended downward. At first, consolidation was mostly reducing the number of small unit banks, but by the early nineties the strong move to consolidation also reached the intermediate and large-size banks with branches (hence the decrease in their number). One important caveat with

[1] The number of banking organizations, that is, the number of banks adjusted to treat all banks within the same holding company as a single bank, decreased steadily from more than 12,000 in 1980 to less than 6,000 in 1998. See Rhoades (2000), which also provides an excellent review of the patterns on merger activity during the period.

[2] Berger and Mester (2003) calculate the average Herfindahl index of concentration in local deposit markets and conclude that the increase in market concentration from 1984 to 1997 has been very moderate.

respect to the decreasing trend in unit banking is that before the complete removal of branching regulations in the early nineties, unit banks were often part of the same bank holding company and hence were managed as, essentially, one bank. Furthermore, after 1989, the cross-guarantee provision in FIRREA[3] implied that solvent banks affiliated with a failing bank were liable for the losses associated with the failure (Walter 1996). In this sense, the distinction between unit banking and banks with branches became less meaningful, and a group of sister banks (that is, banks owned by the same holding company) became not only managed as a single bank, but also legally liable for each other's losses. In summary, the move away from unit banking presented in Panel D of Figure 1, if not interpreted with caution, can induce us to overestimate the economic significance of the change involved.

At the same time that the banking industry was under transformation, the development of money markets and mutual funds created new possibilities for firms to finance their investment and for investors to allocate their funds. It is then a natural question to ask whether the role of banks has been losing importance in the U.S. economy (see also Boyd and Gertler 1994). Panel A of Figure 2 shows that banks' total assets as a proportion of nominal gross domestic product have been relatively stable since the beginning of the seventies. If we interpret gross domestic product as a proxy for the size of the economy, then we may say that commercial banks have roughly kept pace with the secular growth in economic activity. Deposits, though, have become less predominant, but the downward trend is very moderate.[4]

On the other hand, of the total liabilities outstanding in the economy, commercial banks hold a smaller proportion now than they did in the early seventies (see the bold line in Panel B of Figure 2). It is interesting to notice, however, that the decreasing trend has been tapering off and that, as a consequence, the proportion of total debt owed to banks is stabilizing at about 15 percent.[5] A similar long-run trend appears when we consider financial assets held by banks as a proportion of the total liabilities owed by the domestic nonfinancial sector (the thin line in Panel B of Figure 2). This ratio is harder to interpret since some of the financial assets held by banks are liabilities of entities that belong to the financial sector (and hence, they are included in the numerator but not in the denominator). But some interesting facts arise from its evaluation. First, we can see that in the last ten years bank assets have been growing faster than the total debt owed by the nonfinancial sector. Second, the amount of debt owed by the financial sector has been growing faster than the

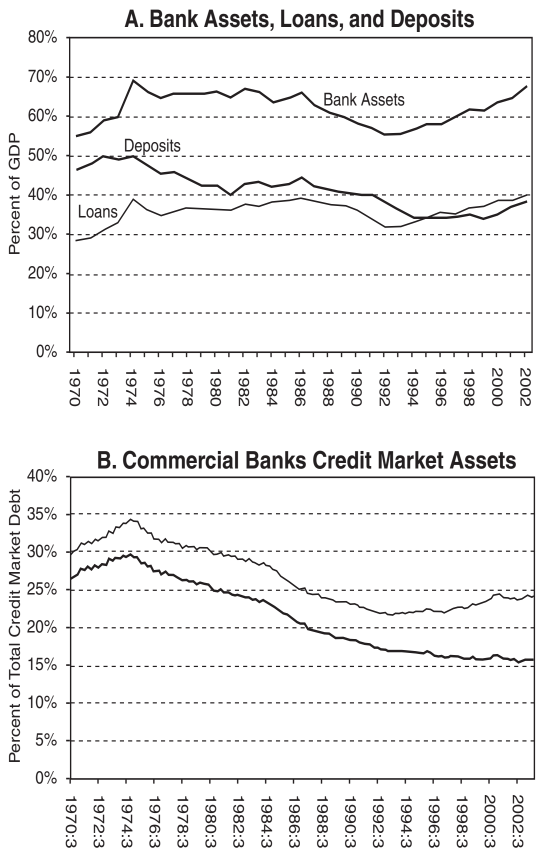
---

[3] FIRREA is the acronym for the Financial Institutions Reform, Recovery, and Enforcement Act of 1989.

[4] The main reason why bank assets have been able to grow at about the same speed of GDP while deposits have not, is that banks have increased their reliance on other borrowed funds to finance their activities (see Figure 3).

[5] Note that by combining the evidence in Panel A and B we can conclude that the total amount of debt outstanding has been increasing faster than nominal GDP. In fact, the ratio of total debt to GDP was 1.75 in 1984, and it is now higher than 3.0.

Figure 2 Banking and the Economy



Note: The bold line in Panel B is the ratio of the total credit market assets held by U.S.-chartered commercial banks to the total credit market debt held by all sectors in the economy. The thin line (top) is the ratio of the total credit market assets held by U.S.-chartered commercial banks to the total credit market debt held by the domestic nonfinancial sector. **Sources:** In Panel A, nominal GDP is from the Bureau of Economic Analysis; bank deposits are from the FDIC-Insured Commercial Banks Table CB15, and bank loans and assets are from the FDIC-Insured Commercial Banks Table CB09. In Panel B, all data are from Table L.1 of the Federal Reserve Board Release Z.1, Flow of Funds Accounts.

amount of debt owed by the nonfinancial sector.[6] In fact, this is true even if

[6] This explains why in the last ten years the thick line in Panel B of Figure 2 has been fairly flat while the thin line has been increasing.

we exclude from the financial sector debt the part that is owed by government-sponsored enterprises (GSE) and other federally related mortgage pools. In other words, financial sector debt has been growing faster, not just at GSEs (a well-known fact), but across the board.

2. THE BALANCE SHEET

During the thirty years under consideration, together with the consolidation trend in banking, the U.S. financial system has experienced several other important changes that, directly or indirectly, influenced the evolution of commercial banks' balance sheets. One prominent example of these developments is the increased participation of mutual funds in financial markets. While the amount of assets held by mutual funds represented less than 1 percent of total financial assets in the economy at the beginning of the 1970s, they now represent more than 11 percent. Since mutual funds are a direct alternative to bank deposits as a channel for savings, their expansion has surely changed the competitive conditions faced by banks in the deposit market. The plan for this section is to review some of the major trends in the evolution of commercial bank balance sheets in order to put in perspective the impact on banking activities of the changes in financial markets occurring in the last thirty years.

Commercial Banks Balance Sheet

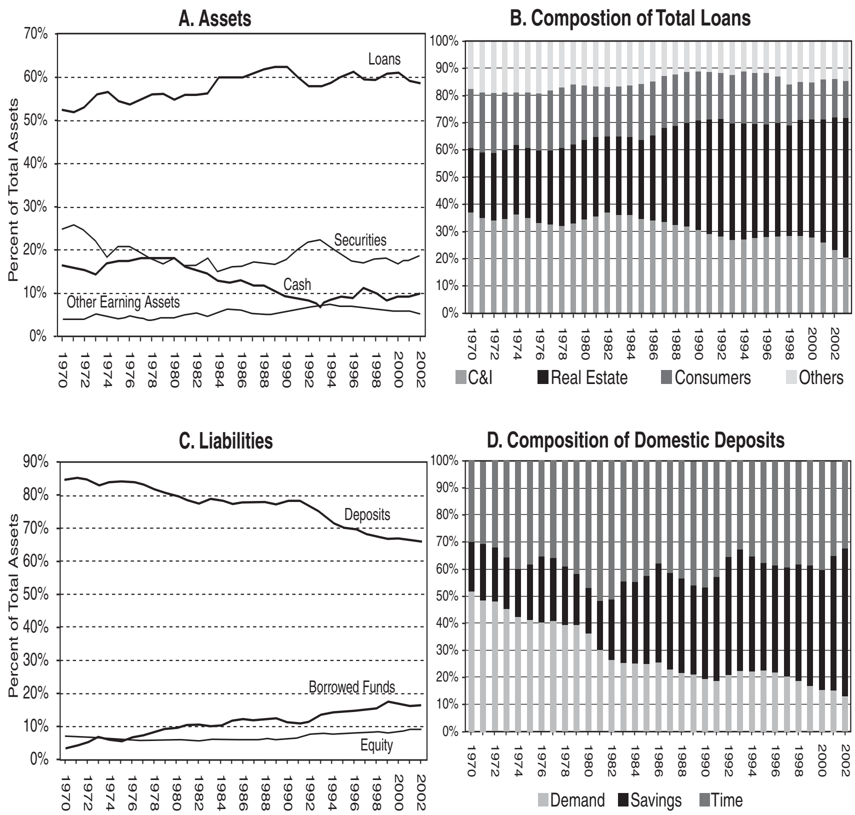
Asset Side	**Liability Side**
Cash	Deposits
Loans and Leases	Borrowed Funds
Securities	Others
Others	Equity Capital

Panel A of Figure 3 shows the long-run trends on the asset side of the U.S. commercial banks' aggregate balance sheet. The proportion of assets represented by loans and leases has stabilized at 60 percent since the early 1980s.[7] Holdings of securities fluctuate around 20 percent, and cash holdings have been trending downwards consistently, partly as a consequence of the implementation of better techniques for cash management.

Panel B of Figure 3 presents the composition of total gross loans (excluding interbank loans) at all commercial banks. We see that real estate loans have been gaining ground while the proportion of commercial and industrial (C&I) loans have been declining for almost twenty years. Since loans are a relatively stable proportion of total assets (see Panel A), the same trends also hold when we look at each type of loan as a proportion of total assets.

[7] During the three decades prior to 1980, the proportion of loans and leases increased from 40 percent to 60 percent while the proportion represented by securities decreased (see Boyd and Gertler 1993).

Figure 3 Trends in Commercial Banks Balance Sheet



Note: Borrowed funds represent federal funds purchased, securities sold under repurchase agreements, demand notes issued to the U.S. Treasury, Federal Home Loan Bank advances, mortgage indebtedness, liabilities under capitalized leases, and all other liabilities for borrowed money. Subordinated notes and all other liabilities constitute a very small (yet growing) portion of the liabilities and hence are not included the figure. **Source:** The data in Panels A, C, and D are from the Federal Deposit Insurance Corporation (Annual Data, Table CB09, CB14). The data in Panel B are from the Federal Reserve Statistical Release (H8, End of Year Data).

It is worth mentioning here that, over the last twenty years, one of the most pronounced changes in the composition of bank loans across the industry is the increasing share of real estate loans in the investment portfolio of small banks (not in the figures). From a level of 22 percent of total assets in 1988, real estate loans at small commercial banks steadily increased their share to a level of more than 37 percent at the beginning of 2003.

Another interesting change in the composition of loans is the strong decrease in the C&I loans' share of total loans during the last three years (see Panel B of Figure 3). This decline is mostly explained by the corresponding decline occurring at large banks (those banks with more than $300 million in assets) where the proportion of C&I loans over total assets went from 19 percent in mid-2000 to around 12 percent in July 2003.

Figure 2 in the previous section suggested that banking has not declined in importance relative to the aggregate economy, or at least not very much. One related question would be to ask whether banking has lost ground as a source of funding for business undertakings. The decreasing contribution of C&I loans to the loan portfolio of banks could be a symptom of such a trend. For a more direct assessment of this question, I constructed Figure 4. Panel A provides the evolution of bank loans to businesses (including not only C&I loans, but also real estate-backed loans to businesses and others) as a proportion of total bank assets. We see that there has been a significant drop in this proportion in the last ten to fifteen years. In principle, this decrease could be the consequence of a slowdown in business activities (relative to other activities in the economy). However, Panel B shows that, during the same period, bank loans have also been losing ground as a source of funds for the nonfinancial corporate business sector. In summary, bank loans to business are becoming less important both for banks and for businesses.

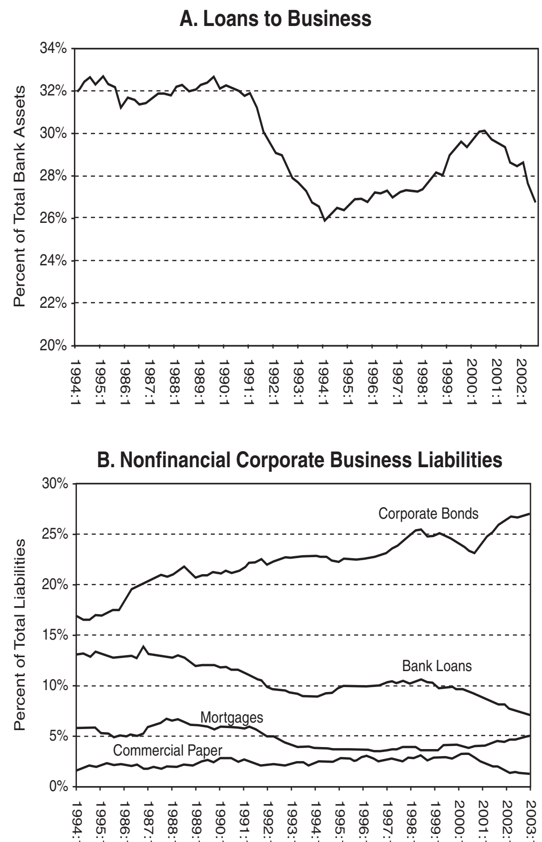
Another important observation coming from Panel A of Figure 4 is that in the last couple of years there has been a steep decrease of the proportion of funds that banks loan to businesses. Even though real estate-backed loans to business (as a proportion of total bank assets) have actually been increasing during the last five years, this increase has not been large enough to offset the pronounced decrease in the contribution of C&I loans in the last two or three years (shown in Panel B of Figure 3). Panel B of Figure 4 shows that, since 1999, bank loans have also been losing share in the total liabilities of nonfinancial corporate businesses.

It is interesting to compare these recent developments with those taking place at the beginning of the 1990s. The decline in bank loans to businesses was very stark during the first three years of the last decade. This decline could be part of the motivation for the "credit-crunch concerns" that were expressed at that time (see Green and Oh 1991). Considering that both 1991 and 2001 were recession years, it seems clear that loans to businesses have shown a strong procyclicality in the last twenty years. This pattern is consistent with the idea of a lending cycle. Yet, it should be clarified that more evidence would be necessary to argue that such a lending cycle is a matter of concern.[8]

---

[8] Weinberg (1995) clearly explains why, even in an efficient credit market, the intensity of loan activity in banks would fluctuate with aggregate conditions. In this sense, to the extent that

Figure 4 Liabilities of Nonfinancial Corporate Business



Note: Loans to business are the result of aggregating all real estate-backed business loans, all commercial and industrial loans, and all agricultural production loans. **Source:** Panel A was constructed using FDIC quarterly data, and Panel B was constructed using data from the Federal Reserve Board Release Z.1, Flow of Funds Accounts.

Going back to Figure 3, on the liability side of commercial banks balance sheets, we see in Panel C that deposits have been losing ground to borrowed funds as a source of funding for banks. The increased importance of mutual funds and of money market instruments are an important part of the explanation for this trend. To the extent that investment in mutual funds has become a close substitute for deposits, banks are less able to rely on the latter to finance their

business investment has been especially weak during the 2001–2002 period, it should not come as a surprise that bank loans to business have also shown relative weakness during this period.

lending activities. Also, banks are using more money market instruments to manage their short-run liquidity needs, and such liabilities are included in the "borrowed funds" category. Finally, in the last ten years, there has been an increase in the amount of loans that commercial banks take from the Federal Home Loan Banks system. These loans are also part of the borrowed funds aggregate.

Panel D shows that the proportion of deposits represented by time deposits has been relatively stable over the years, but the proportion of demand deposits has been decreasing steadily, from being more than 50 percent of total deposits in 1970 to less than 15 percent in 2003.[9] Panel D also shows that in the last ten years, savings accounts have been crowding out both time and demand deposits.[10]

During the nineties, the share of total bank assets funded by core deposits (total deposits less time deposits that are larger than $100,000) has been decreasing even faster than the share of total deposits (see Genay 2000).[11] Banks have increased their reliance on more interest-sensitive liabilities for funding their loan activities. In fact, the loan-to-deposit ratio is now above one (from about 0.6 at the beginning of the seventies), implying that banks are loaning out a higher amount of funds than they have in deposits.[12] While borrowed funds are more sensitive to interest rate changes, they also require lower reserve levels to satisfy unexpected swings in the demand for liquidity. The lower need for reserves increases the propensity of banks to loan available funds and hence increases the loan-to-deposit ratio.

One of the most important components of the liability side of banks' balance sheets is, of course, equity capital. Table 1 shows some interesting patterns in the holdings of equity capital by U.S. commercial banks.

First, smaller banks tend to hold higher capital ratios (defined as equity capital over total assets). This evidence is consistent with the view that smaller banks tend to be less diversified and need to hold relatively more equity capital

---

[9] Since the 1980s a type of transaction account called Negotiable Order of Withdrawal (NOW) has become more widely used. NOW accounts are not included in the demand deposit category used in Panel D of Figure 4. It is interesting to note, however, that their increase in importance has not been enough to compensate for the ground lost by demand deposits (another type of transaction account). By dividing deposits into transaction and nontransaction accounts we can observe that in the seventies, transaction accounts were around 50 percent of the total, but, since then, they have been steadily losing participation to become less than 20 percent in 2002.

[10] This increase in the importance of savings accounts versus demand deposits accounts is partly associated with the increase in the use of "sweeping" since 1994. Sweep accounts allow banks to periodically reclassify balances from retail transactions deposits into savings accounts so as to reduce their reserve requirements (see Krainer 2001 for details).

[11] This decreasing trend in core deposits has been reversed in the last couple of years, a pattern consistent with the fact that this has been a period of low interest rates and especially weak performance of the stock market (see Carlson and Perli 2002).

[12] The loan-to-deposit ratio is calculated as the ratio of total gross loans and leases and total domestic deposits in all FDIC insured commercial banks.

Table 1 Equity Capital (Percent of Total Assets)

Size	<$100 Million	$100 Million– $ 1 Billion	$ 1 Billion– $10 Billion	>$10 Billion
1985	8.50	7.20	5.84	4.91
1990	8.98	7.67	6.33	5.26
1995	10.42	9.39	8.57	7.19
2000	11.06	9.59	8.98	8.07
2003*	11.27	9.94	10.61	8.52

Note: (*) The data for 2003 are only for the first half of the year. **Source:** FDIC Quarterly Banking Profile.

to control the impact of agency costs in their access to external financing (see, for example, Ennis 2001).

Second, we can see that capital ratios have been increasing through time for all bank sizes. During the eighties and nineties important changes in bank regulation have increased the role of capital as a way of limiting the risk exposure of banks. The Basel Accord in 1988 is, without a doubt, a clear move toward increasing bank holdings of capital.[13] In fact, the biggest jump in capital ratios at commercial banks occurred during the first four years of the nineties when both the capital-to-asset ratio (in Table 1) and the risk-adjusted capital ratio increased around 2 to 3 percent, on average. The risk-adjusted capital ratio is calculated using a risk-adjusted asset base on which each component of the bank's assets is weighted according to a regulatory-risk classification. It should be said, however, that the risk classification is relatively coarse and hence the adjustment may fail to accurately reflect the evolution of risk in banks' portfolios. Still, to the extent that the adjustment does take into account, at least partially, the changes in risk, it is interesting to note that, since 1992, the risk-adjusted capital ratio has been relatively constant at around 12.5 percent. This evidence, combined with the information in Table 1, suggests that asset portfolios of banks have become riskier during the last ten years.[14]

The risk-adjusted capital ratio started to be calculated officially only after 1990, and hence it is not possible to take a long-run perspective on its evolution. The simpler captial-to-assets ratio presented in Table 1 allows us to obtain a longer-run perspective on the evolution of capital (since 1985). Even though

[13] The "prompt corrective action" clause introduced in the FDICIA uses Basel capital requirements to determine when the regulators should intervene in the case of an undercapitalized bank. As a consequence, the importance of capital ratios as a supervisory tool has increased since 1991.

[14] The measure of risk-adjusted assets contains a component representing off-balance-sheet exposure of banks, which has been gaining importance during the last decade.

this simpler ratio fails to account for changes in assets' risk, it can still be informative (as a complement to the risk-adjusted ratio) to the extent that banks exercise regulatory-capital arbitrage to artificially reduce the measure of adjusted assets. In support of this view, for example, Estrella, Park, and Peristiani (2000) argue in favor of using simple capital ratios as informative signals for predicting bank failures.

3. THE INCOME STATEMENT

I now turn to reviewing the long-run trends in the components of the income statement of banks. Obviously, the changes in the structure of the industry (see Section 1) and in the composition of the balance sheet (see Section 2) can change the sources of income at banks. Another recent development in banking that had significant implications for the evolution of the income statement is the increase in importance of off-balance-sheet activities. Some of the most common off-balance-sheet activities are the provision of lines of credit, the securitization and sale of loans, and the trading of derivative instruments. Outstanding lines of credit, for example, are an important source of fee income for banks. Some of these activities have become common practice only relatively recently, and the data that would allow for the kind of long-run perspective that I wish to provide in this article are not readily available. For this reason, I will not explicitly cover this aspect of the evolution of banking in the recent past, but it should be kept in mind when interpreting the trends in the income statement that I will discuss next.[15]

There are five main components of the income statement of commercial banks: interest income and expense, non-interest income and expense, and loan loss provisions. Interest income is the result of all interest and dividends earned by banks on interest-bearing assets (such as loans and leases). Interest expense is the result of all interest paid to depositors and other creditors of the bank. Net interest income, then, is the difference between interest income and interest expense. Non-interest income includes fee income, gains on securities transactions, and all other income not originated in interest payments. Non-interest expense includes personnel compensation, legal expenses, office occupancy and equipment expense, and other expenses. Finally, provision for loan losses is the amount charged as operating expenses to provide an adequate reserve to cover anticipated losses in the loan portfolio. These charges become part of the allowance for loan losses, a negative component on the asset side of the banks balance sheet, which is then used to charge off loans after they become nonperforming.

[15] For a thorough discussion of commercial banks' off-balance-sheet activities, see Boyd and Gertler (1993, 1994).

Income Statement

- Net Income
- Net Interest Income
- Interest Income
- Interest Expense (–)
- Provisions for Loan Losses (–)
- Non-interest Income
- Non-interest Expense (–)

Note: The negative sign in parenthesis (-) indicates that the component is subtracted when computing net income.

In the last ten years, the annual return on assets (ROA) at commercial banks in the United States has been historically high: the average since 1993 has been 1.15 percent, while from 1950 to 1985 the average was 0.72 percent.[16] There are several possible factors that could explain this sustained change in level. One possibility is that there has been an increase in the efficiency with which banks manage their assets. (This change constitutes a movement from the interior of the set of feasible risk-return combinations toward the frontier of that set.)[17] A second possibility is that banks are engaging in riskier (on- and off-balance-sheet) activities that are associated with a higher rate of return. (This change is a movement along the risk-return frontier of the feasible set.)

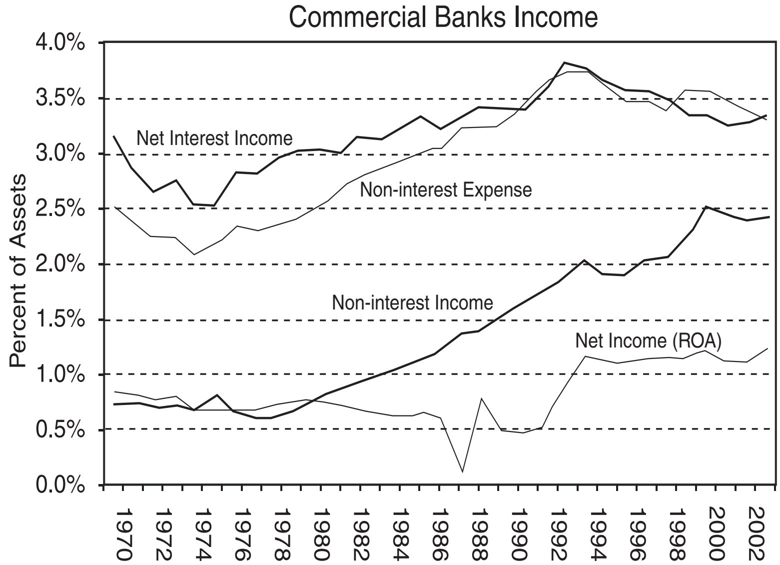
The evidence seems to indicate that the explanation for this increase in return must be a combination of the two possible causes. On the one hand, banks may be moving toward increasing efficiency. In Figure 5 we can see that since the beginning of the nineties, coinciding with the increase in ROA, the ratio of non-interest expenses to total assets has been decreasing steadily. The decrease of total operational expenses relative to assets could be indicative of an improvement in the efficiency level with which banks handle their assets. Another way of thinking about aggregate movements toward the risk-return frontier is to think that an increasing number of institutions are getting closer to implementing the current best managerial practices for banks. Berger and Mester (2003) provide some evidence that seems to indicate that the average inability of banks to adopt best practices has decreased during the 1990s (see their Table 3). On this same line of inquiry, Berger and Humphrey (1992) undertake a detailed analysis of the 1980s, a period of rising inefficiency due to dispersion inside the best-practice frontier.

On the other hand, banks may be taking more risks. As we saw in the previous section, the risk-adjusted assets base used for regulatory purposes has

[16] Annual return on assets is defined as the ratio of annual net income and total average assets. Since the banking industry experienced a generalized crisis from 1986 to 1992, those years were left out for the sake of long-run trend comparison.

[17] Inefficiency in banking may originate not only in banks having higher resource cost than necessary, but also in the fact that some banks may be holding a portfolio of assets that, conditional upon risk, is dominated in rate of return by other feasible portfolios.

Figure 5 Income Statement



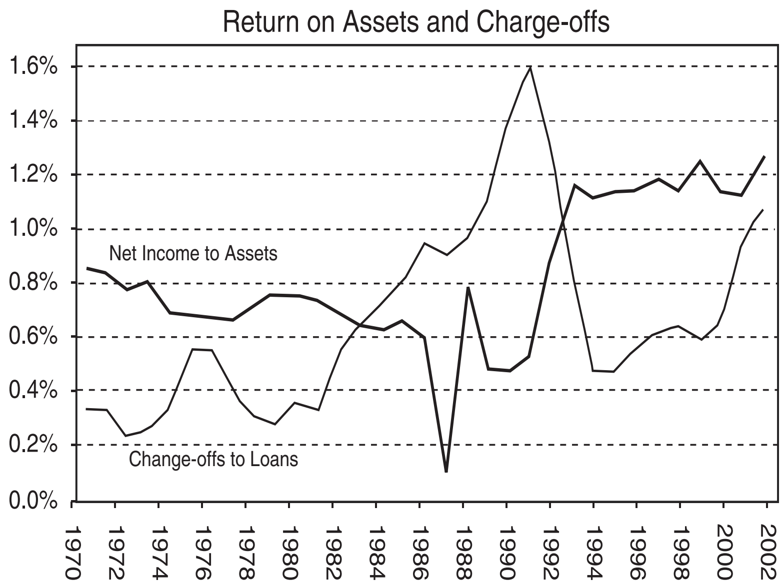
Source: Annual data from the FDIC-Insured Commercial Banks Table CB04.

been growing at a higher rate than unadjusted bank assets during the nineties. This fact implies that under the (imperfect) regulatory risk classification of assets, banks are indeed taking on more risk. Another piece of evidence that supports this view is given in Figure 6, which shows both the evolution of return on assets and charge-offs at commercial banks since 1970. We can see in the figure that, together with the increase in return on assets since 1993, the level of charge-offs as a proportion of loans remained relatively high compared with its historic level of around 0.3 percent. If we think that the charge-offs to loans ratio is positively correlated with the risk of the loan portfolio, then we can conclude that since 1993 both risk and return have increased in the banking industry.[18]

A third possible explanation for the increase in the return on banks' activities is based on the notion that the 1990s was a period of repeated innovations in banking (such as credit scoring, widespread ATM networks, and many others). Early adopters of new technologies tend to earn supernormal profits for

[18] Berger and Mester (2003) use the average standard deviation of annual return on assets as a proxy for bank risk. They show that this indicator has actually been decreasing during the nineties, but, due to lack of data, they do not present estimates for the seventies and early eighties.

Figure 6 The Risk-Return Frontier



Source: Annual data from the FDIC-Insured Commercial Banks Table CB04 and Table CB08.

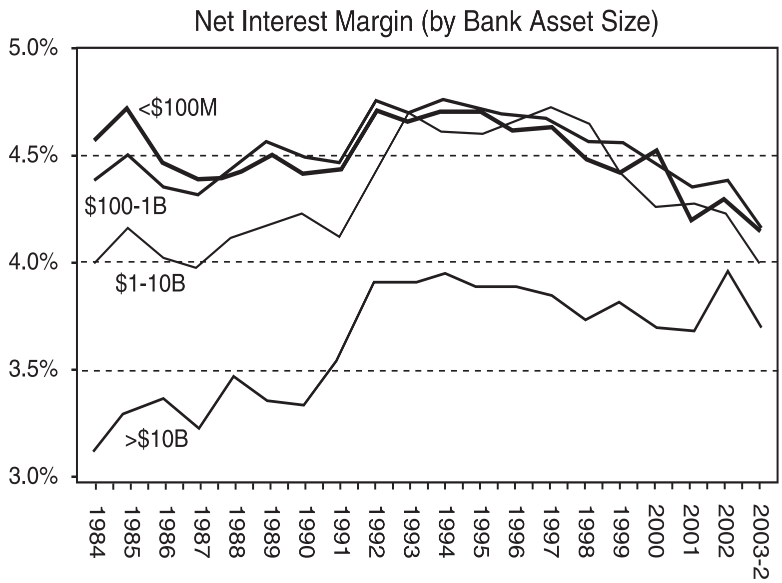
some time until the technology becomes widely used by competitors. Repeated innovations can then explain a long period of high returns like the one the banking industry experienced during the 1990s. Berger and Mester (2003) suggest this possibility but do not provide direct evidence supporting the hypothesis.

In general, the evolution of the return on assets has been relatively uniform across bank sizes. In particular, ROA decreased for all bank sizes during the second half of the eighties and the beginning of the nineties and then recovered to levels of above 1 percent after 1993. The exception to this uniformity of behavior across sizes is the large variation in ROA during 1987–88 (see Figure 5). These variations at the industry level mostly reflect the changes in ROA at banks with more than $10 billion in assets (the largest category).[19]

With respect to the evolution of charge-offs across bank sizes, the story is somewhat different. While small banks (those with less than $100 million in

[19] In 1987, provision for loan losses at large banks spiked up significantly (creating the opposite effect on ROA). This increase in provisions is mostly explained by the need to write off large international loans to less developed countries that became nonperforming after the international debt crisis of the eighties.

Figure 7 Interest Income at Commercial Banks

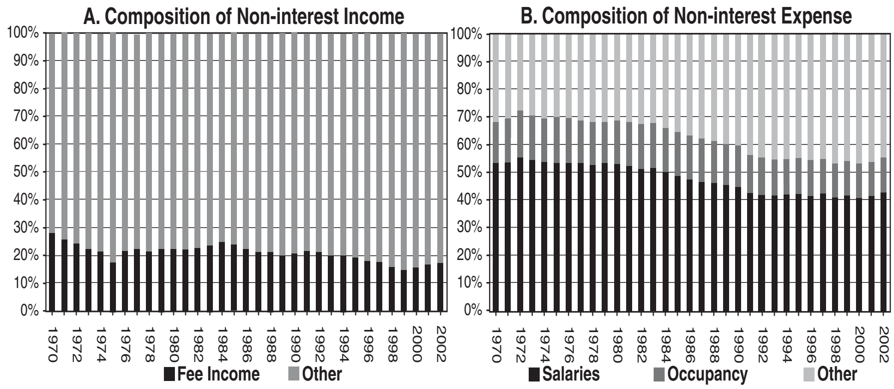


Note: The net interest margin is the ratio of net interest income and interest earning assets. The data are for four different sizes of banks: banks with more that $10 billion in assets, banks with $1 billion to $10 billion in assets, banks with $100 million to $1 billion in assets, and banks with less than $100 million in assets. **Source:** FDIC Quarterly Banking Profile. The data for 2003 are only for the first half of the year.

assets) had relatively high levels of charge-offs during the mid-eighties, the spike in the late eighties and early nineties (see Figure 6) is fully explained by the increase in charge-offs at medium-to-large banks (those with more than $1 billion in assets). Similarly, the increase in charge-offs during 2001 is mostly concentrated in these medium-to-large size banks. Overall, charge-offs tend to be fairly procyclical, but, again, the behavior across different sizes of banks is not very uniform. In particular, during the 2001 recession, the level of the charge-off ratio at banks with less than $1 billion in assets does not show any significant increase.

The net interest margin is one of the most common indicators of profitability in traditional banking activities (that is, holding deposits and lending). This indicator results from expressing net interest income as a percentage of (average) interest-earning assets. Figure 7 shows that medium-to-large banks experience a significant increase in the net interest margin at the beginning of the nineties. On average, net interest margins have been higher at all banks

Figure 8 Non-interest Income and Expense



Note: Fee income represents the amount of all service charges on deposit accounts in domestic offices such as maintenance fees, activity charges, administrative charges, overdraft charges, and check certification charges. Other non-interest income includes: income from fiduciary activities; gains, losses, and fees relating to foreign currency or foreign exchange transactions; gains, losses, and fees from assets held in trading accounts; net gains from the sale or disposition of loans, premises (including branches and offices) and fixed assets, and other real estate owned; all service charges, fees, and commissions (other than those relating to deposits in domestic offices); fees charged on bank-issued credit cards; net gains on futures and forward contracts; and other miscellaneous income.
Source: Annual data from the FDIC-Insured Commercial Banks Table CB07.

during the nineties. These higher margins are consistent with an increase in default and interest-rate risk at banks.[20] Angbazo (1997) provides further evidence on this link between interest margins and risks at banks. He also shows that off-balance-sheet activities tend to increase risk and interest rate margins at banks.

Figure 8 shows the trends in the composition of non-interest income and expenses. While non-interest income has increased substantially during the last twenty years (see Figure 5), its composition has not changed much (see Panel A of Figure 8). In particular, income from fees on deposit accounts has been stable at around 20 percent of total non-interest income for a long time. In summary, it is true that in the last twenty years fee income has been increasing relative to total gross income at banks, but other non-interest

[20] Interest rate risk is the exposure of a bank's financial condition to adverse movements in interest rates (see Basle Committee 1997).

income has also become increasingly important.[21] Boyd and Gertler (1994) suggest that the evolution of the relative contribution of other non-interest income to total income is a good proxy for the increase in importance of off-balance-sheet activities at banks. In this respect, after growing steadily for the last twenty years, other non-interest income has gone from representing less than 10 percent of bank income in the seventies to representing more than 25 percent of such income today.

With respect to non-interest expense, there is a clear trend toward a lower contribution of salaries and employee benefits to total non-interest expenses. This decrease is almost exactly matched by the increase in the proportion of other expenses, while the contribution of occupancy expenses remains constant. In the last couple of decades there has been a tendency for banks to outsource employment-intensive activities to other affiliates of the bank holding company or to service bureaus (Berger and Mester 2003). The cost of these outsourced activities becomes part of the "other" component of non-interest expense at the same time that it reduces the "salaries" portion of total expense.

4. CONCLUSION

The banking trends reviewed in this article suggest some conclusions. First, even though commercial banking activities have been changing significantly, banks as a group are still a very important player in the financial markets of the U.S. economy. Second, banks do seem to be moving away from their traditional activities of handling deposits and providing loans to business, but this trend is fairly gradual. Third, while banking activities have become more profitable in general, the evidence suggests that they have also become riskier.

Finally, as I suggested in the introduction, taking a long-run perspective can help to identify some patterns of the response of modern banking to economic slowdowns. Although a careful study of this issue was beyond the scope of this article (see Schuermann 2004), from the comparison of the data from 1990–91 and 2001, some regularities can be identified. For example, bank loans to businesses seem to be very procyclical, and both charge-offs and net interest margins seem to be relatively countercyclical.

---

[21] For a detailed study of the evolution of banks' retail fees during the nineties, see Hannan (2001).

REFERENCES

Angbazo, Lazarus. 1997. “Commercial Bank Net Interest Margins, Default Risk, Interest-Rate Risk, and Off-Balance Sheet Banking.” *Journal of Banking and Finance* 21: 55–87.

Basel Committee on Banking Supervision. 1997. “Principles for the Management of Interest Rate Risk.” *Basel Committee Publications* 29 (September).

Berger, Allen N., and David B. Humphrey. 1992. “Measurement and Efficiency Issues in Commercial Banking.” In *Output Measurement in the Service Sectors*. Ed. Zvi Griliches. Chicago: University of Chicago Press: 245–79.

__________, and Loretta J. Mester. 2003. “Explaining the Dramatic Changes in Performance of U.S. Banks: Technological Change, Deregulation, and Dynamic Changes in Competition.” *Journal of Financial Intermediation* 12: 57–95.

Broaddus, J. Alfred, Jr. 1998. “The Bank Merger Wave: Causes and Consequences.” Federal Reserve Bank of Richmond *Economic Quarterly* 84 (Summer): 1–11.

Boyd, John H., and Mark Gertler. 1993. “U.S. Commercial Banking: Trends, Cycles, and Policy.” National Bureau of Economic Research *Macroeconomic Annual*: 319–68.

__________. 1994. “Are Banks Dead? Or Are the Reports Greatly Exaggerated?” Federal Reserve Bank of Minneapolis *Quarterly Review* 18 (Summer): 2–23.

Carlson, Mark, and Roberto Perli. 2003. “Profits and Balance Sheet Developments at U.S. Commercial Banks in 2002.” *Federal Reserve Bulletin* (June): 243–70.

Ennis, Huberto M. 2001. “On the Size Distribution of Banks.” Federal Reserve Bank of Richmond *Economic Quarterly* 87 (Fall): 1–25.

Estrella, Arturo, Sangkyun Park, and Stavros Peristiani. 2000. “Capital Ratios as Predictors of Bank Failures.” Federal Reserve Bank of New York *Economic Policy Review* (July): 33–52.

Genay, Hesna. 2000. “Recent Trends in Deposit and Loan Growth: Implications for Small and Large Banks.” *Chicago Fed Letter* 160 (December).

Green, Edward J., and Soo Nam Oh. 1991. “Can a ‘Credit Crunch’ Be Efficient?” Federal Reserve Bank of Minneapolis *Quarterly Review* 15 (Fall): 3–16.

Hannan, Timothy H. 2001. “Retail Fees of Depository Institutions, 1994–99.” *Federal Reserve Bulletin* (January): 1–11.

Krainer, John. 2001. “Retail Sweeps and Reserves.” Federal Reserve Bank of San Francisco *Economic Letter* (January).

Rhoades, Stephen A. 2000. “Bank Mergers and Banking Structure in the United States, 1980–98.” Board of Governors of the Federal Reserve System Staff Study 174 (August).

Schuermann, Til. 2004. “Why Were Banks Better Off in the 2001 Recession?” Federal Reserve Bank of New York *Current Issues in Economics and Finance* 10 (January).

Walter, John R. 1996. “Firewalls.” Federal Reserve Bank of Richmond *Economic Quarterly* 82 (Fall): 15–39.

Weinberg, John A. 1995. “Cycles in Lending Standards?” Federal Reserve Bank of Richmond *Economic Quarterly* 81 (Summer): 1–18.

Which Price Index Should a Central Bank Employ?

Roy H. Webb

In the 1970s the United States experienced inflation rates that were high relative to any other decade in the nation's peacetime experience. During that decade the consumer price index doubled, rising at a 7.4 percent average annual rate. At one point in the early 1980s, the CPI inflation rate exceeded 14 percent for a full year. When inflation was that high, the choice of which price index to employ to calculate inflation was a secondary concern for policymakers. As Figures 1 and 2 will indicate later in this article, commonly used price indexes gave the same message: inflation in the 1970s and early 1980s was relatively high.

The situation now is different. At low rates of inflation, differences among price indexes become more important. While it is difficult to imagine the difference between 10.0 and 10.5 percent inflation affecting monetary policy, the difference between 1.0 and 1.5 percent inflation could lead to different policy choices. Yet different price indexes can easily yield inflation rates that differ by that 50 basis-point magnitude. Thus in this period of low inflation, the choice of which price index to use has become an important issue for monetary policy analysis.

This article begins with the premise that a central bank places a high weight on keeping inflation low. Several central banks have adopted a formal inflation target by making a public commitment to achieving a particular goal for inflation, as discussed in Bernanke et al. (1999), for example. Central banks in other countries, including the United States, while not setting formal inflation targets, have nonetheless made it clear that low inflation is an important policy concern.

The author gratefully acknowledges helpful comments from Huberto Ennis, Marvin Goodfriend, Andreas Hornstein, and Thomas Humphrey. The views and opinions expressed in this article are solely those of the author and should not be attributed to any other person or the Federal Reserve Bank of Richmond or the Federal Reserve System.

Based on that premise, this article explores several considerations that lead to the choice of which price index to employ for setting monetary policy. Several widely used price indexes are discussed, and the author presents evidence that favors one particular index.

1. WHICH PRICE INDEX?

First Choices

Several grounds are given that could be used to choose which price index to employ. As this article progresses, the set of possible choices will be narrowed until one remains.

Credibility

A central bank seeking to maintain low inflation must at some point acquire credibility for being willing and able to take the actions necessary to achieve its goals. As part of a strategy for low inflation, then, that central bank must also employ a price index that itself has credibility. That is, the price index should be the result of a well-grounded statistical program that is not subject to political manipulation. The United States has many credible price indexes produced by the Bureau of Labor Statistics and the Bureau of Economic Analysis.

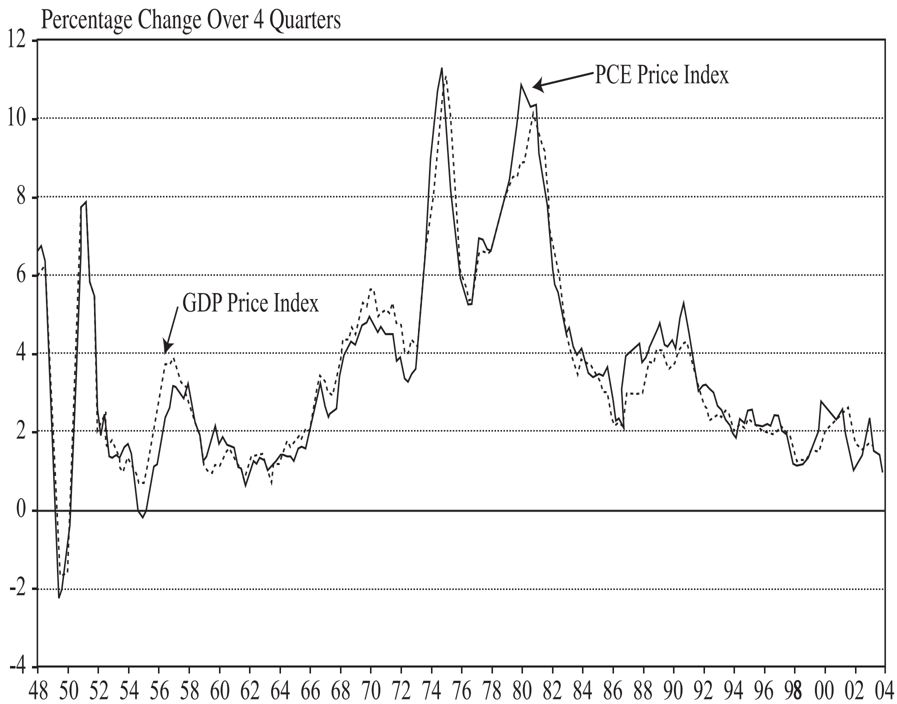
Breadth

The next choice is between a narrow price index, which includes prices of only a few items, or a broad price index with many items. Some analysts have advocated a narrow index of prices of a few raw materials on the grounds that those prices can respond rapidly to changes in monetary conditions. Well-known examples include commodity price indexes that have been published by the Commodity Research Bureau and the *Journal of Commerce*. An important drawback is that those prices can also respond rapidly to supply shifts of individual items, and as a result, movements in the index can reflect relative price changes rather than general price changes. Thus central banks have long given more prominence to broad price indexes in their policy deliberations and have chosen broad price indexes for inflation targets.

Sector

A wide variety of broad price indexes are published, including producer price indexes, price indexes for GDP and its components, and consumer price indexes. Looking at the major broadly based indexes, it is clear that they are highly correlated. Figure 1 shows inflation rates from two indexes that cover different sets of prices. The GDP price index covers goods and services produced in the United States, whereas the price index for personal consumption

Figure 1 Inflation Rates



expenditure (PCEPI) covers consumer spending in the GDP accounts. Since inflation rates calculated from those indexes are very similar, the choice can be based on the need to acquire and maintain credibility with the public. It is probable that members of the public are more likely to accept a monetary strategy for low inflation if they can relate it to their everyday experience. Thus a measure of consumer prices that is believed to be relevant to individual households would be a natural choice. Consequently, every central bank that has an explicit inflation target has chosen a measure of consumer prices.

2. THE CHOICE BETWEEN TWO CONSUMER INDEXES

In the United States, the best-known measure of consumer prices is the Consumer Price Index (CPI)[1] published by the Bureau of Labor Statistics. It has a

[1] Actually two versions of the CPI are published. The CPI-U covers all urban consumers, whereas the CPI-W covers urban wage and salary workers. In practice, the two indexes give virtually identical inflation rates, and thus the two will not be distinguished in the text. The figures in this article include the CPI-U.

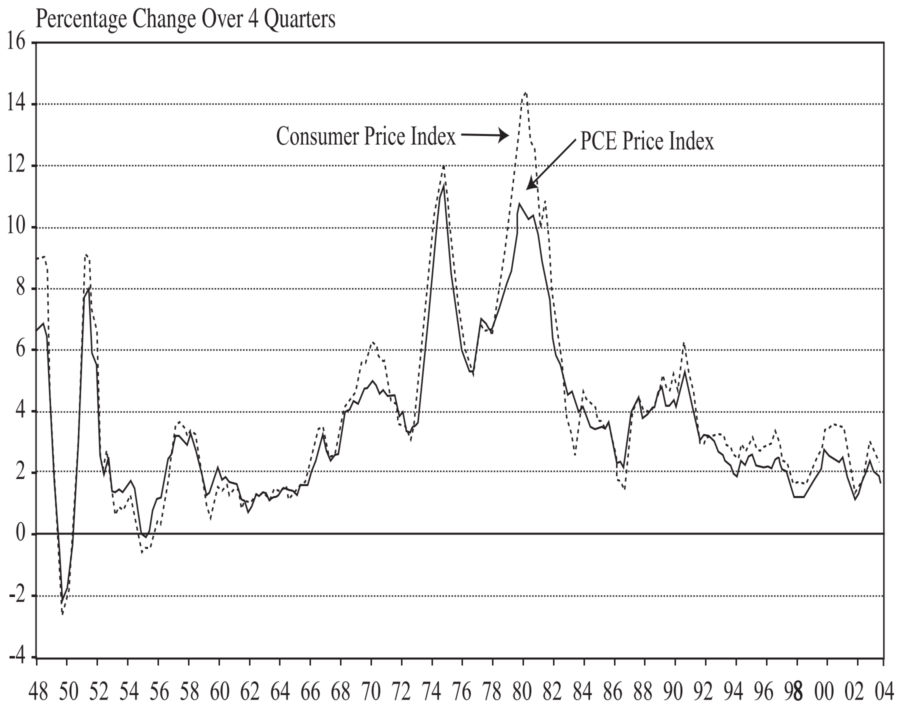
long track record and is widely used as an inflation index in government spending and taxing programs as well as in private contracts. The CPI's credibility has been enhanced by efforts of its producers to make a wealth of technical information readily available to the public on the details of constructing the index. The Bureau of Labor Statistics conducts an active research program that has helped the index adapt to changes in the economy and improve over time.

Setting the index apart from similar indexes in most other countries, the Commissioner of Labor Statistics has made a public commitment to using economic theory to guide important decisions that are made in constructing the index (Abraham 1997). Specifically, the concept of a cost-of-living index is now used as an organizing principle for making decisions concerning the production of the CPI. A cost-of-living index can be defined as the minimum expenditure required in a particular period to attain the same standard of living as was achieved in a reference period, divided by actual expenditure in the reference period. Economic theory tells how a cost-of-living index can be calculated from a consumer's preferences (for example, Diewert 1987), and the resulting index will correctly convert nominal income to real income. Statistical agencies in other countries have apparently shied away from employing cost-of-living methodology because it can be difficult to apply in real-world situations. The alternative, though, is that indexes constructed without that discipline can be hard to interpret. For example, the price of owner-occupied housing is the largest single component, by far, in the CPI; yet the price of owner-occupied housing is totally omitted in consumer price indexes in several other countries. That omission could not be defended in a cost-of-living framework.[2]

Although the cost-of-living concept helps its producers answer practical questions that arise as the index is produced, the CPI is not a cost-of-living index. The Bureau of Labor Statistics did not scrap the existing CPI when it decided to use the cost-of-living index as a benchmark. Instead, there have been incremental improvements to the index since then. In comparing the current CPI with an ideal cost-of-living index, the single most important difference is the formula that is used to construct the CPI. That formula does not account for the possibility of consumers responding to changing relative prices by changing their expenditure patterns. Later in this article there will be a more detailed discussion of the CPI's formula, and the appendix contains a

[2] The perennial question of whether to include asset prices in the CPI can be evaluated in the context of a cost-of-living index. That approach indicates that the price of *services* of consumer durables should be in the index and that there are two valid approaches to estimating the services of consumer durables. One is a Jorgenson (1963) user-cost formula, and the other estimates an imputed flow of services. The latter is currently used in the CPI for owner-occupied housing, which estimates an owner's equivalent rent from rental prices of similar structures. Importantly, the consumption of the services of a durable asset is independent of the method of financing the asset's purchase. That financing decision is thus outside the scope of a cost-of-living index.

Figure 2 Consumer Inflation Rates



numerical example that may help illustrate why the CPI is not a cost-of-living index.

The CPI thus has many positive attributes. If it were the only index of consumer prices available, it could be the basis for a successful monetary strategy aiming for low inflation. However, another index has some advantages over the CPI. The PCEPI attempts to cover the prices of all items consumed by residents of the United States. As Figure 2 indicates, while broad movements in the two indexes are similar, at times the differences have been substantial.

One source for the differences in Figure 2 is the changing methodology that has been used in constructing the CPI. The PCEPI uses a consistent methodology for its entire history; whenever that methodology has been changed, past numbers were accordingly revised. But values of the CPI are not changed after being published.[3] Thus in the late 1970s and early 1980s, housing prices

[3] Since the CPI is widely used to index money payments, fixing previously published values of the index avoids the question of whether payments that were previously made would need to be recalculated each time a methodological change was made to the CPI that altered historical values.

were overstated in the CPI, and inflation rates calculated using the CPI for that period were consequently overstated (Blinder 1980).[4]

More relevant for current monetary policy, there are other differences between the two indexes that affect current values of the indexes. Most important is that while both indexes are weighted averages of prices, two different formulas are used to calculate those averages. The CPI, a Laspeyres index, uses weights for individual prices that represent an item's importance in consumer expenditure at a fixed point in time (in 2003, the weights were based on average spending in the years 1999 and 2000). In symbols, the exact formula is

$$CPI_t = \frac{\sum_i q_{i,b} p_{i,t}}{\sum_i q_{i,b} p_{i,b}}, \tag{1}$$

where CPI_t is the value of the consumer price index at time t, $q_{i,b}$ is the quantity of item i consumed in the base period b ($b \neq t$), $p_{i,t}$ is the price of item i in period t, and $p_{i,b}$ is the price of item i in the base period. In contrast, the PCEPI is a Fisher Ideal index, the geometric average of a Laspeyres index like the CPI and an index that uses current values of spending for the weights on prices. In symbols,

$$PCEPI_t = \sqrt{\frac{\sum_i q_{i,t-1} p_{i,t}}{\sum_i q_{i,t-1} p_{i,t-1}} \frac{\sum_i q_{i,t} p_{i,t}}{\sum_i q_{i,t} p_{i,t-1}}}, \tag{2}$$

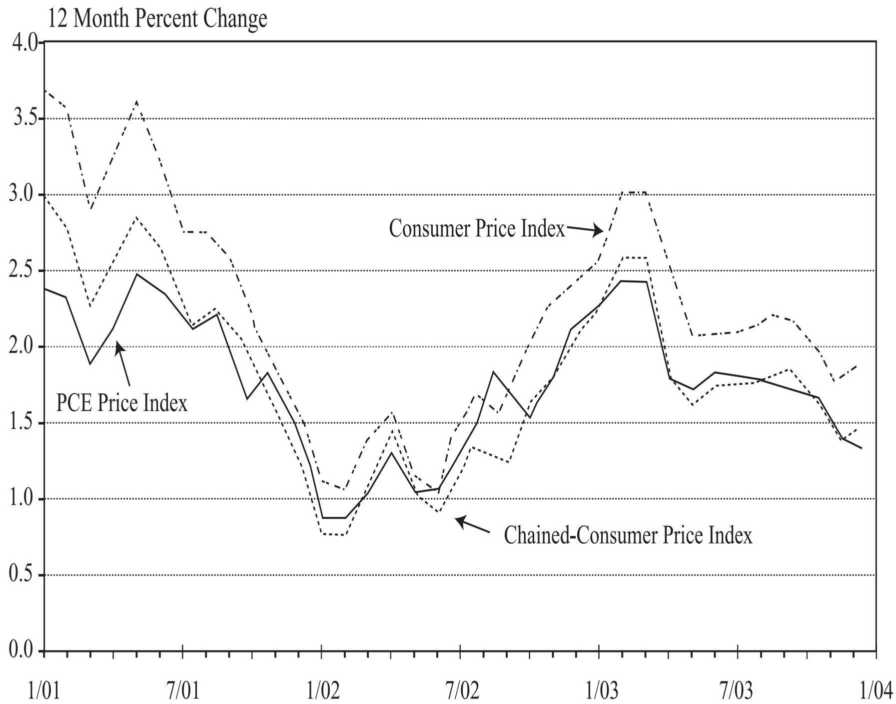
where $q_{i,t}$ is the quantity of item i consumed in period t. Note that the formula for the PCEPI includes data on current period quantities $q_{i,t}$ that are omitted from the formula for the CPI. Also, the PCEPI does not have a fixed base period; instead, the index values are calculated using data from the current period and the previous period. Since there is not a designated base period, the index number for one particular period will be set to 100. The formula in (2) is then used to link adjacent periods together.

The difference in formulas is important because the CPI does not routinely allow for changing expenditure patterns in response to relative price changes. This could be particularly important when technical progress results in falling prices of goods such as computers, cellular phones, and television sets. Failure to account for increasing spending on items with falling prices would create a bias in the index that would lead it to rise more rapidly than the true cost of living. The Fisher Ideal index, however, allows for changing expenditure patterns in a manner that allows it to approximate a cost-of-living index especially well and is thus known as a *superlative* index (Diewert 1987).

---

[4] At that time, the CPI attempted to measure an average price of items purchased by a representative consumer, not the cost of living. It therefore included the purchase price of owner-occupied housing plus a measure of mortgage interest rates. By the mid-1980s, after the damage was done and the index had overstated inflation, the current approach of pricing the service flow from owner-occupied housing was adopted.

Figure 3 Consumer Inflation Rates



The appendix illustrates the construction of a Laspeyres index and a Fisher index in a simple case with a large change in the pattern of expenditure.

It may seem that the difference in formulas would be an example of esoteric trivia; however, at low inflation rates, the magnitude of the difference can be large when compared with the absolute rate of inflation. In order to focus on the difference between a superlative index and a Laspeyres index, it is helpful to consider briefly a new index from the Bureau of Labor Statistics. The Chained-CPI (C-CPI) uses exactly the same price information as the CPI but is based on another type of superlative index, a Tornquist index.[5] Like the Fisher index, the Tornquist index includes information on current period quantities and thereby allows for changing expenditure patterns. Figure 3 shows inflation rates calculated using the CPI, the C-CPI, and the PCEPI.

[5] The formula for the Tornquist Index is $\prod_i (\frac{p_{i,t}}{p_{i,t-1}})^{(\frac{s_{i,t-1}+s_{i,t}}{2})}$ where s_i is the expenditure share of item i, that is, $\frac{q_i p_i}{\sum_j q_j p_j}$. Note that current expenditures enter the formula through the expenditure share term.

There is a noticeable difference between the CPI and the other two indexes. Over this period (the entire period for which the C-CPI is available), the average difference between the CPI and C-CPI was 44 basis points, which is entirely attributable to the different formulas. The average difference between the two superlative indexes was only 5 basis points.

In addition to the different aggregating formulas, the prices and relative importance of various items can differ between the CPI and PCEPI. Most of the individual prices in the PCEPI are identical to those in the CPI. The most notable exception is spending for medical services, where the PCEPI uses information from producer price indexes. Moreover, a few items are covered in one index but not another. Also, the relative importance of a particular item can differ considerably between the two indexes, since completely different sources of information are used to determine relative importance. The PCEPI uses information to construct GDP, such as economic census data and industry trade data. The CPI uses information from periodic Consumer Expenditure Surveys. Some analysts (such as Lebow and Rudd 2001) have viewed the weights in the PCEPI as likely to be more accurate. In the survey data used for the CPI, a member of a household is asked to give information on spending of all members of the household. If items accounting for a small portion of spending tend to be missed or forgotten in the household survey, then the fraction of spending for big-ticket items would tend to be biased upward. Not surprisingly, then, owner-occupied housing has a much larger weight in the CPI than in the PCEPI. To quantify the effect of different weights, Lebow and Rudd compared the published CPI with an alternative CPI using PCE weights. From 1987 to 2000, the average inflation rate was 10 basis points lower when using the PCE weights.

Accordingly, due to the clearly superior formula for computing the index and the probably superior item weights, changes in the PCEPI should provide a better estimate of the true cost of living.

Limitations of Price Indexes

It is important to consider some limitations of both indexes. In a dynamic economy, the items available for purchase are constantly changing, with new items being introduced continuously, some old items being improved, and other old items disappearing from the market. Accounting for new items is a challenge for producers of price indexes. For example, the Boskin Commission Report (1996) noted that although there were 36 million cellular phones in use at the time of the report, there was no price of cell phones in the CPI. Compounding the problem is the typical product cycle, in which a new good initially sells for a relatively high price, but as economies of scale are realized and new competitors enter the market, the price falls rapidly before eventually leveling out. If the price of a new item does not enter a price index promptly, then the interval of a rapidly falling price can be missed entirely, and the price index would therefore overstate inflation. Since the Boskin report was written, the Bureau of Labor Statistics has reduced—but not eliminated—the time it takes for a new item to enter the CPI.

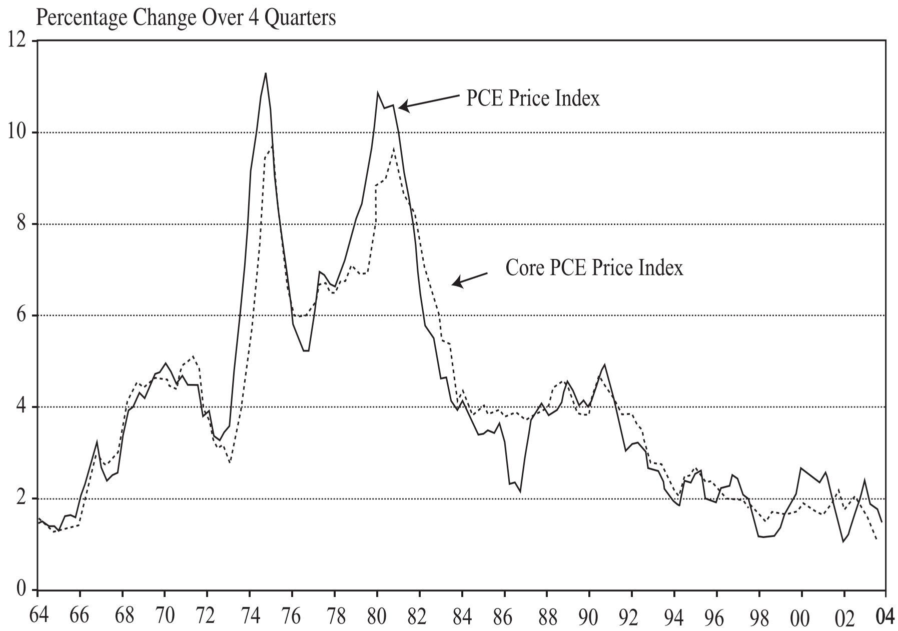
A related difficulty is accounting for quality change. If the greater durability or improved functionality of a product is not taken into account, then a price index will overstate the cost of living. And if quality improvements routinely outweigh quality deterioration, inflation can be overstated. Some detailed studies of particular products have found that accounting for quality change would have made a sizeable difference in recorded prices. Based on many of these studies, the Boskin Commission Report estimated that there was a 60 basis-point upward bias in CPI inflation rates at that time, due to new products and quality change. A more recent estimate by Lebow and Rudd puts the bias at 37 basis points. It should be emphasized that these estimates are subject to a large amount of imprecision. If it were easy for analysts to disentangle the portion of price changes that reflect quality change, it would probably be part of routine price index calculation already. At this time, improving estimates of quality change is an ongoing challenge for statistical agencies.

3. CORE INDEXES

A final choice is between the PCEPI and a measure of *core* inflation. For purposes of monetary policy, it would not be desirable to respond to temporary changes in measured inflation that are likely to be reversed. Thus policymakers in many countries pay particular attention to a core price index that excludes some items that account for a significant amount of short-run volatility in the index but do not have much effect on the long-run trend. For several decades, inflation analysts in the United States have focused on a core price index that excludes food and energy prices. As illustrated in Figure 4, the core PCEPI is less volatile than the overall index; using the core index reduces the variance of inflation rates in that figure by 31 percent. Most importantly, the core index omits some significant fluctuations in the overall index that were soon reversed but could have led to inappropriate monetary policy actions. For example, the core index did not decline significantly in 1986 and did not rise significantly in 2002–2003. In 1986, crude oil prices fell sharply and led to lower retail energy prices. In the latter episode, as energy prices increased sharply in response to the approach of war in Iraq, the overall index signaled rising inflation, but the core index signaled low, falling inflation. Finally, removal of energy and food prices has had a small effect on the long-run trend. Over the 40-year period illustrated in Figure 4, the PCEPI increased at a 4.05 percent annual rate, whereas the core PCE index increased at a 3.95 percent rate.

Most of the countries with full inflation targets have taken a similar approach and employed a core index that omits a few items. Besides food and

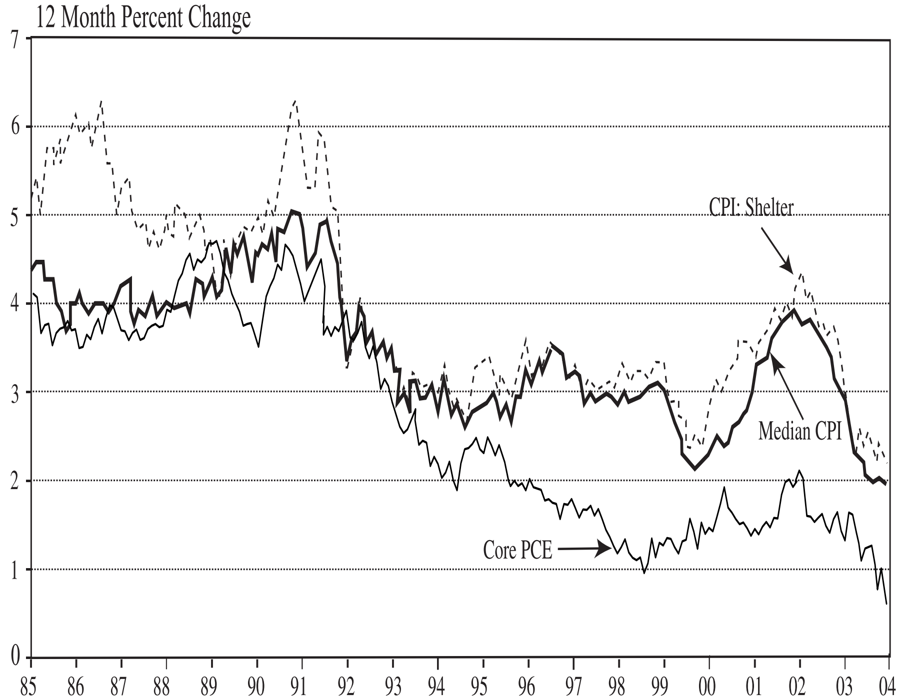
Figure 4 Core Inflation



energy prices, several countries omit indirect taxes. These widely used core indexes were not derived from economic or statistical theory, but were instead based on the judgment that their use would result in better choices of monetary policy actions. Researchers have also examined alternatives that use more elaborate statistical methods for determining core inflation (see, for example, the survey by Johnson 1999). At this point, though, this is only research in progress that has not resulted in a new standard for determining core inflation. The major hurdle for these studies will be to demonstrate that using a proposed method would lead to better monetary policy decisions.

This point can be illustrated with a particular alternative estimate of core inflation. A *median* CPI (Bryan and Cecchetti 1993) is based on the statistical property that a median is not influenced by extreme observations, unlike the arithmetic average used for the CPI and the PCEPI. The Federal Reserve Bank of Cleveland accordingly calculates a median CPI as an alternative measure of core inflation and posts recent and historical values on their web site. Despite that prominence, however, their median CPI has not supplanted the traditional core index. Two observations may explain why. First, due to the large weight

Figure 5 FRB Cleveland's Median CPI, Core PCE Price Index, and CPI: Shelter



placed on housing expenditure in the CPI, the median CPI often simply picks up the behavior of housing prices, as is illustrated in Figure 5. Thus the correlation between the 12-month change in the shelter component of the CPI and the median CPI was 0.90 since 1985 and was even higher before then. In contrast, the correlation between the nonshelter component of the CPI and the median CPI was only 0.48 since 1985. Also, note that from January 2000 to November 2001, the inflation rate calculated from the median CPI *increased* from 2.4 to 4.0 percent, which at face value would indicate an excessively easy monetary policy stance and might signal the need to raise the federal funds rate target. Over that period, however, the economy weakened in 2000 and moved into recession in 2001. Thus it appears that at that time the traditional core price indexes gave a better guide for monetary policy. Accordingly, the 12-month change in the core PCEPI remained below 2.2 percent in 2000 and 2001.

Although the traditional core index has not been supplanted by an alternative, there is a strong case for continued research on alternative measures of core inflation. Given the imperfect nature of macroeconomic statistics, one should always wonder if any particular statistic is giving misleading signals, and the core PCEPI is no exception. Having well-studied alternatives could thus be valuable to policymakers if at any time the traditional core index were to be in doubt.

4. CONCLUSION

This article studies which price index to use for determining monetary policy actions. The best choice for the United States is currently the core price index for personal consumption expenditure. From 1996 to the end of 2003, the four-quarter change in that index remained within a narrow range, 0.9 to 2.1 percent. Like other macroeconomic statistics, price indexes are not precision tools. Allowing for about a half percentage point of upward bias in the reported inflation rate, the true cost of living has been rising very slowly for several years.

APPENDIX

A numerical example may help clarify the effects of different index formulas. Assume that there are two goods, apples and bananas, denoted with superscripts a and b. There are two time periods, 0 and 1. Money income is y, and utility is u. I assume, $y_0 = \$12.00$, $p_0^a = p_0^b = \$1$ and $u = \sqrt{q^a q^b}$, where p represents a price and q is a quantity. Given the initial conditions, utility is maximized at a level of 6.0 when 6 apples and 6 bananas are consumed. Using the equations in the text, both a Laspeyres index L and a Fisher index F will have values of 1.00 in period 0.

Now, let $y_1 = \$12.00$, $p_1^a = \$1.50$, $p_1^b = \$0.50$, and we can ask if real income has risen, fallen, or remained unchanged. If we divide money income by a cost-of-living index, the real income rises if and only if utility rises. Given the utility function and new prices, the optimal quantities are $q_1^a = 4$ and $q_1^b = 12$, and utility rises to approximately 6.93. In other words, the price changes have allowed utility to increase significantly once quantities are allowed to adjust.

Consider first the Laspeyres formula given in equation 1; substituting the values above gives $L_1 = \frac{6\times1.5+6\times0.5}{6\times1+6\times1} = 1.00$, and, therefore, real income in period 1 is \$12.00. Although utility rose, real income, calculated by using a Laspeyres index, did not change. Now use the Fisher formula of equation 2: $F_1 = \sqrt{\frac{6\times1.5+6\times0.5}{6\times1+6\times1} \times \frac{4\times1.5+12\times0.5}{4\times1+12\times1}} = \sqrt{\frac{3}{4}}$, and real income is approximately \$13.86. Thus the latter index correctly leads to rising real income with rising utility.

Finally, for comparison we can construct a cost-of-living index. We can define the value as exactly 1.00 in period 0. The utility function has the property that utility is maximized when exactly half of income is spent on each item. Using that knowledge, the minimum expenditure in period 1 that achieves the utility level of 6 is $3\sqrt{12}$, or approximately \$10.39, which purchases approximately 3.46 apples and 10.39 bananas. Thus the cost-of-living index for period 1 is $\frac{3\sqrt{12}}{12}$, which in this case is exactly equal to the Fisher index. In both cases, a utility-maximizing consumer would buy more bananas, which became less expensive and, correspondingly, fewer apples, which became more expensive. Both the cost-of-living index and the Fisher index correctly captured that changing expenditure pattern.

REFERENCES

Abraham, Katharine G. 1997. Testimony before the Subcommittee on Human Resources, House Committee on Government Reform and Oversight, April 30. Available at http://www.bls.gov/news.release/cpi.br043097.brief.htm (accessed April 6, 2004).

Bernanke, Ben S., Thomas Laubach, Frederic S. Mishkin, and Adam S. Posen. 1999. *Inflation Targeting: Lessons from the International Experience.* Princeton: Princeton University Press.

Blinder, Alan. 1980. "The Consumer Price Index and the Measurement of Recent Inflation." *Brookings Papers on Economic Activity* 2: 539–65.

Boskin, Michael J., Ellen Dullberger, Robert J. Gordon, Zvi Griliches, and Dale Jorgenson. 1996. "Toward a More Accurate Measure of the Cost of Living." Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index, December 4. Available at http://www.ssa.gov/history/reports/boskinrpt.html (accessed April 6, 2004).

Bryan, Michael F., and Stephen G. Cecchetti. 1993. "Measuring Core Inflation." Federal Reserve Bank of Cleveland Working Paper 9304.

Diewert, W. E. 1987. "Index Numbers." *The New Palgrave: A Dictionary of Economics*. Ed. John Eatwell, Murray Milgate, and Peter Newman. New York: Stockton Press 2: 767–80.

Johnson, Marianne. 1999. "Core Inflation: A Measure of Inflation for Policy Purposes." Bank for International Settlements Conference Paper.

Available at http://www.bis.org/publ/bisp05_p4.pdf (accessed April 6, 2004).

Jorgenson, Dale W. 1963. "Capitol Theory and Investment Behavior." *American Economic Review* 53 (May): 247–59.

Lebow, David E., and Jeremy B. Rudd. 2001. "Measurement Error in The Consumer Price Index: Where Do We Stand?" Federal Reserve Board of Governors, Finance and Economics Discussion Series, 2001-61.