http://clevelandfed.org/research/review/
1989 Q 1
Best available copy

ECONOMIC REVIEW

1989 Quarter 1
Vol. 25, No. 1

The Determinants of Direct Air Fares to Cleveland: How Competitive? **2**
by Paul W. Bauer and Thomas J. Zlatoper

A spate of recent fare increases has focused attention on the airline industry's competitiveness. Many analysts are concerned about continuing changes in the industry in the wake of its consolidation in the last few years. The authors examine the competitiveness issue by developing a model for the determinants of air fares. They find that, controlling for other important factors, air fares are lower on routes that have more carriers offering service, although most of the decrease in fares is achieved once two or three carriers offer service.

Bank Capital Requirements and the Riskiness of Banks: A Review **10**
by William P. Osterberg and James B. Thomson

Previous studies of the impact of capital requirements on bank portfolio decisions typically assume that the deposit rate paid by banks is not a function of the riskiness of the bank's portfolio. Such studies conclude that stiffer capital requirements decrease portfolio risk, but may increase the probability of bankruptcy. The authors show that the variance of earnings and the incentive to increase leverage are reduced with risk- and leverage-related deposit rates. However, the impact of increased capital requirements on portfolio behavior is generally ambiguous.

Turnover, Wages, and Adverse Selection **18**
by Charles T. Carlstrom

Empirical studies of turnover and wages have shown that frequent job-changers have lower average wages and flatter age-earnings profiles than workers who change jobs infrequently. This paper argues that adverse selection in the labor market can explain this phenomenon. Lower-productivity workers are the frequent job-changers, according to Akerlof's "lemon" principle, giving rise to the observed relationship between turnover and wages. Adverse selection also provides a basis for examining the welfare implications of low-productivity workers in the labor market.

Economic Review is published quarterly by the Research Department of the Federal Reserve Bank of Cleveland. Copies of the *Review* are available through our Public Affairs and Bank Relations Department, 216/579-2157.

Coordinating Economist: Randall W. Eberts

Editors: William G. Murmann
Robin Ratliff
Design: Michael Galka
Typesetting: Liz Hanna

Opinions stated in *Economic Review* are those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

Material may be reprinted provided that the source is credited. Please send copies of reprinted material to the editor.

ISSN 0013-0281

The Determinants of Direct Air Fares to Cleveland: How Competitive?

by Paul W. Bauer
and Thomas J. Zlatoper

Paul W. Bauer is an economist at the Federal Reserve Bank of Cleveland. Thomas J. Zlatoper is a professor at John Carroll University and a research associate at the Center for Regional Economic Issues at Case Western Reserve University. The authors would like to thank John R. Swinton and Paula J. Loboda for their expert research assistance, and thank Randall W. Eberts for his comments.

Introduction

Eleven years ago, Congress decided in the form of the Airline Deregulation Act of 1978 that the operational decisions of airlines—where planes can fly and what fares can be charged—would be better left to the airlines than to the regulators.

This decision has caused numerous changes in the industry: discount fares have become widespread and traffic has boomed, new carriers have come and gone, hub-and-spoke networks have emerged, and frequent-flier plans have become the rage. As long as the industry remains competitive, many analysts assert that travelers have little to fear from these continuing changes, since competition ensures that fares are held close to cost and that economically viable service is provided.

With the consolidation of the airline industry that started in 1986, many analysts have begun to wonder about its competitiveness, both now and in the future. The wave of mergers has resulted in an increase in the number of airlines that offer nationwide service, but this comes in the form of "fortress hubs." At such airports, the dominant carrier typically offers about three-quarters of the airport's flights. In addition, the national carriers now face less competition from regional and local service carriers, many of whom have been purchased by or signed operating agreements with the national carriers. The impact of these developments (and of possible future consolidations) on fares depends on the competitiveness of the markets for air travel.

To gain insight into the competitiveness of the airline industry, this paper examines the determinants of air fares for first-class, coach, and discount service to a particular destination: Cleveland, Ohio. We begin by examining two of the market models that have been proposed for the airline industry. The first is the traditional view that market competitiveness is determined by the number and concentration of firms in the market. The second is the theory of contestable markets, in which the number of actual competitors in the market plays only a small role. According to this theory, it is the number of carriers that could potentially enter the market that constrains fares.

We then discuss the implications for appropriate public policy. A reduced-form equation for air fares is constructed, and the data that were collected to estimate its parameters are described. Finally, we present and analyze the empirical results and discuss the implications for public policy.

Our results suggest that these markets (the airline routes) are not perfectly contestable. The number of actual competitors does influence the fares charged by the airlines, other things being equal. Thus, policymakers should act where possible to ease entry barriers in the industry in order to preserve and enhance competition.

I. Economic Models of Airline Competition

The traditional method of determining the amount of competition in a market is to examine the market shares of the largest firms operating in that market. This measure is relevant because, until recently, most economists thought that competitiveness was determined by the number and concentration of the actual participants in the market.

The U.S. Department of Justice uses a Hirschman-Herfindahl Index (HHI: the sum of the squares of all of the firms' market shares) as an aid in assessing the impact of proposed mergers on market competition. This index ranges from close to zero in the case of a perfectly competitive market to 10,000 (100^2) in the case of a monopoly.[1] The Department of Justice guidelines recommend rejecting mergers that result in markets with an HHI greater than 1,800 unless the resulting increase in the HHI is less than 50 or there are some other special considerations. The rationale is that fewer competitors reduce the competitiveness of the market, since there will be less pressure to hold down prices and costs and since the firms will find it easier to collude.

The airline industry appears to be very uncompetitive when one examines the HHIs of various airline routes. According to a recent Congressional Budget Office study, on a typical route only 2.5 carriers offer service. Even if these carriers each had an equal share of the market, this would result in an HHI of over 4,000.

The U.S. Department of Transportation—the agency charged with oversight of the airline industry—has taken a different approach than the Justice Department. Over the last few years, it has allowed mergers to occur between carriers even when many of their routes overlapped. For example, TWA and Ozark competed on many routes involving their joint hub of St. Louis, and their merger in 1986 resulted in a large increase in concentration on these routes. In 1983, the HHI was about 3,100; just after the merger, the HHI was about 5,800; and in 1988 the HHI had risen to about 6,800, with TWA offering about 82 percent of the flights out of St. Louis. The TWA-Ozark merger was clearly outside the Department of Justice's guidelines discussed above (however, there was the special consideration that Ozark was in financial difficulty and might have failed unless it was taken over).

In approving mergers such as this one, the Department of Transportation relied heavily on the relatively new theory of contestable markets developed primarily by Baumol, Panzar, and Willig (1982).[2] This theory states that under certain conditions, it is not necessary to have a large number of firms actually operating in a market in order for prices and output in that market to approximate the ideal outcome of a perfectly competitive market. If entry barriers into the market are low, and if there are no irrecoverable costs to exiting the market, then even markets with only a few firms will be constrained to follow the same marginal-cost pricing that perfect competition with many firms would. If the firms in the market tried to raise prices above marginal cost (the extra cost of producing an additional unit of output), then entrepreneurs could enter the market and charge a slightly lower price than the incumbent firms (taking away those firms' customers) and could earn an above-average profit. The ease of entry and exit from a perfectly contestable industry means that potential competitors also exercise competitive pressure on the firms in the industry.

There were several reasons to believe that the airline industry might approximate a perfectly contestable market after the Civil Aeronautics Board stopped regulating routes and fares, a process phased in over several years starting in the late 1970s. Planes now can quickly be shifted from one route to another, and many of the airlines rent a significant proportion of their aircraft fleets. In addition, there is a ready secondary market for used aircraft, so a major component of an airline's capital stock is much easier to acquire and dispose of than in most other industries.

Working against the idea that the airline industry is perfectly contestable are the current congestion problems in the air traffic control network. Also, new entrants find it difficult to acquire gate space and slots for takeoffs and landings at the more congested airports. Computer reservation systems, travel agent commissions, frequent-flier plans, and hub-and-spoke

■ 1 Since the market shares are squared before summing, the market shares of the largest firms will influence the index the most.

■ 2 The theory of contestable markets has been applied to a number of other industries. Whalen (1988) finds evidence that the banking industry is perfectly contestable.

networks are also cited as characteristics of providing air service that make entry into new markets difficult. Borenstein (1988) provides a more detailed investigation of these issues.

If the market for air fares approximates a perfectly competitive market, then there is very little need for government oversight of the economic conditions in the airline industry, although there still would be a role in the regulation of air safety. Actual and potential competitors force the airlines serving a market to provide the service that passengers want at the lowest possible fares. If the market is not perfectly contestable, then the government can ensure that entry into the market is as free as possible, and should enforce existing antitrust laws to protect consumers by preserving as much competition in the market as possible.

II. Empirical Model and Data

Although other researchers (for example, Bailey, Graham, and Kaplan [1985], Borenstein [1988], Butler and Huston [1987], and Call and Keeler [1985]) have explored the extent of competition in the airline industry by using models similar to the one we develop, none of these studies employs data as recent as ours (April 1987). Thus, not only are our data further away from the beginning of deregulation, but they also follow the latest wave of mergers that occurred in 1986.

The following observations will be useful in constructing the testable hypotheses. If the market were perfectly contestable, then the number of carriers serving a route would have no relationship to passenger fares. If potential competitors constrain the fare-setting abilities of existing carriers, then the market is imperfectly contestable and the effect of the number of carriers serving a route should have a significant, although small, effect on the fares charged. Lastly, if entry is so blocked that existing carriers have little to fear from new entrants, then the degree of competition on a route will be determined by the number of carriers currently serving the route, and the effect of an additional carrier on the route could cause a significant reduction in fares. This is the more traditional view of the relationship between the degree of competition and the number of competitors.

In comparing the fares charged with the number of carriers on the route across routes, one must allow for other factors that influence fares. In essence, we are estimating a reduced-form equation for air fares, so that anything that influences the demand for, or the cost of, air travel should be taken into account. The most important of these factors are the length of the route, the volume of traffic on the route, and whether one or both of the airports involved are hubs or are restricted in takeoff and landing slots.

The characteristics of a particular flight on a given route can also influence both the supply and the demand for the flight. The most important of these are the number of stops on a particular flight, whether a meal is provided, and the particular carrier offering the flight. Finally, the demand for air service on a particular route will depend in part on characteristics of the flight's origin and destination cities, such as their average per capita incomes and whether they are business or tourist centers.

We estimate the following model using ordinary least squares (OLS):

$$
\begin{aligned}
(1)\quad \text{FARE} = {} & a_0 + a_1\ \text{CARRIERS} \\
& + a_2\ \text{CARRIERS}^2 + a_3\ \text{PASS} \\
& + a_4\ \text{MILES} + a_5\ \text{MILES}^2 \\
& + a_6\ \text{POP} + a_7\ \text{INC} + a_8\ \text{CORP} \\
& + a_9\ \text{SLOT} + a_{10}\ \text{STOP} \\
& + a_{11}\ \text{MEAL} + a_{12}\ \text{HUB} \\
& + a_{13}\ \text{EA} + a_{14}\ \text{CO} + \text{error},
\end{aligned}
$$

where
- FARE = one-way air fare;
- CARRIERS = number of carriers;
- CARRIERS^2 = number of carriers squared;
- PASS = total number of passengers flown on route (all carriers);
- MILES = mileage from the origin city to Cleveland;
- MILES^2 = the number of miles squared;
- POP = population of the origin city;
- INC = per capita income of the origin city;
- CORP = proxy for potential business traffic from the origin city;
- SLOT = dummy variable equaling 1 if the origin city has a slot-restricted airport, 0 otherwise;
- STOP = number of on-flight stops;
- MEAL = dummy variable equaling 1 if a meal is served, 0 otherwise;
- HUB = dummy variable equaling 1 if the origin city has a hub airline, 0 otherwise;

EA = dummy variable equaling 1 if the carrier is Eastern Airlines, 0 otherwise;
CO = dummy variable equaling 1 if the carrier is Continental Airlines, 0 otherwise.

This model is estimated separately for each of three classes of fares: first class, coach, and restricted discount.

The data to estimate this model were combined from a number of sources. The *Official Airline Guide* (April 1987) was the source of the fare information and the data on the flight characteristics, such as CARRIERS, STOP, SLOT, MEAL, EA, and CO. All of the data pertain to direct domestic flights terminating in Cleveland. Unfortunately, fares for connecting flights could not be analyzed here because only direct fares are reported in the *Official Airline Guide.* In future research, we hope to obtain such data.

Data on passengers (PASS) and nonstop mileage from origin to destination (MILES) were taken from the U.S. Department of Transportation's *Origin and Destination City Pair Summary.* Data on per capita income (INC) of the origin cities were obtained from the *Survey of Current Business* (April 1986 issue). The number of Standard & Poor's companies headquartered in each origin city (CORP) was compiled to be used as a proxy for the business traffic likely to be generated by each city. Information on whether an origin city was considered to have a hub airline (HUB) was obtained from 1985 Department of Transportation statistics. For each of the three fare classes, summary statistics on the variables used in the analysis are provided in table 1.

III. Estimation Results

Tables 2, 3, and 4 report OLS estimates of equation (1) for first-class, coach, and discount fares. The amount of variation in fares explained in each estimated equation (the adjusted R-square statistics in tables 2 through 4) is generally high, and is higher for the first-class and coach categories than for the discount category. This is probably the result of the discount fares being less homogeneous than the other fare classes. For our discount fare, we always selected the least expensive restricted-discount fare reported in the *Official Airline Guide,* and these fares were not always subject to exactly the same restrictions.[3]

In interpreting these results, recall that only direct flights to Cleveland were included in the data. Also note that since more than 90 percent of passengers travel on some type of discount fare,

TABLE 1

Summary Statistics of the Variables

	First-Class Fares		Coach Fares		Discount Fares	
Variable	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
FARE	330.17	123.63	201.78	89.60	62.65	29.85
CARRIERS	2.77	1.33	2.89	1.25	2.88	1.25
PASSENGERS	18,458.00	22,802.00	15,260.00	21,414.00	15,273.00	21,406.00
MILES	744.19	535.18	537.27	465.43	541.25	466.32
INCOME	13,996.00	1,766.00	13,709.00	1,643.60	13,727.00	1,656.10
CORP	10.63	16.67	8.76	15.17	8.75	15.17
SLOT	0.22	0.42	0.19	0.39	0.19	0.39
STOP	0.46	0.60	0.41	0.63	0.42	0.63
MEAL	0.60	0.49	0.44	0.50	0.44	0.50
HUB	0.71	0.46	0.66	0.47	0.66	0.47
CO	0.16	0.37	0.08	0.27	0.08	0.27
EA	0.16	0.37	0.07	0.26	0.08	0.27
POP	4,046.30	4,668.20	3,497.60	4,184.90	3,493.40	4,187.80

SOURCE: Authors' calculations.

■ 3 It was not possible to select one particular type of discount fare for all of the routes because no type of discount fares were reported for all routes.

TABLE 2

First-Class Fare Estimates

Variable	Estimated Coefficient	Standard Error	T-Ratio
CARRIERS	-19.50	22.20	-0.878
CARRIERS^2	2.79	4.42	0.632
MILES	0.233	0.455E-1	5.13
MILES^2	-0.974E-5	0.197E-4	-0.495
POP	-0.598E-2	0.357E-2	-1.67
INC	-0.195E-2	0.285E-2	-0.686
CORP	3.62	1.05	3.45
PASS	-0.818E-3	0.106E-2	-0.771
STOP	12.50	9.18	1.36
SLOT	7.13	23.90	0.299
HUB	11.30	12.60	0.900
MEAL	11.20	10.50	1.07
EA	-18.30	11.40	-1.60
CO	-66.40	11.60	-5.72
CONSTANT	212.00	40.60	5.21

NOTE: All values are authors' calculations. Number of observations = 163; R-squared = 0.863.

according to the Air Transport Association, this class of service is probably the most important for evaluating the competitiveness of the industry.[4]

The first issue is the effect of the number of carriers on fares. The estimated values for CARRIERS and CARRIERS^2 have the expected signs for all three classes of fares. These results suggest that as additional carriers begin service on a route, fares are lowered, since CARRIERS is negative. But because the coefficient of CARRIERS^2 is positive, each additional carrier lowers fares on the route less than the one before. After three or four carriers are serving a route, fares no longer appear to be affected by the number of carriers.

These coefficients are statistically significant for coach and discount fares, but are not significant for first-class fares. For discount fares, the addition of one carrier to a monopoly route would lower fares by about $11, other things being equal. Adding a third carrier to the route would again lower fares, but by only about $6.50. With a fourth carrier, fares drop even less, by about $2. Fares do not appear to fall any more once about four carriers are serving the route. At this point, discount fares are about $20 less than they would be if only one carrier served the route. Extrapolation beyond this point is not warranted since the maximum number of carriers on any route in our sample is only five.

The above result for first-class fares does not mean that these fares are perfectly contestable, however. If we estimate the same model as equation (1), but replace CARRIERS and CARRIERS^2 with a dummy variable equal to one if there is more than one carrier on the route and zero otherwise, we find that the coefficient of this variable is significant and negative for first-class fares. First-class fares are about $21 lower on routes with more than one carrier, other things being equal. In other words, since fares are cheaper on routes with more than one carrier, these results do not support the notion that these routes are perfectly contestable.

Earlier studies that investigated whether the market for air fares was perfectly contestable also found little support for perfect contestability. As mentioned above, their data generally came from the early 1980s and thus may have been estimated too soon after deregulation for the airlines to have adjusted to their new environment. Because our study employs fare data from April 1987, it is unlikely that the lack of contestability is a result of the airlines' having insufficient time to adjust to the deregulated environment. This data set also has the advantage of being gathered about a year after the merger wave that peaked in 1986.

Not surprisingly, MILES has a positive and significant estimated coefficient for each class of fares. Coach and discount fares have a significant amount of "fare taper": as the flight distance increases, the cost per mile falls. First-class fares

TABLE 3

Coach Fare Estimates

Variable	Estimated Coefficient	Standard Error	T-Ratio
CARRIERS	-23.00	11.60	-1.99
CARRIERS^2	4.00	2.19	1.83
MILES	0.277	0.231E-1	12.00
MILES^2	-0.520E-4	0.104E-4	-4.98
POP	-0.114E-2	0.200E-2	-0.570
INC	-0.178E-2	0.168E-2	-1.06
CORP	1.22	0.487	2.51
PASS	-0.275E-3	0.522E-3	-0.527
STOP	7.64	3.59	2.13
SLOT	-0.746	11.20	-0.667E-1
HUB	4.18	5.16	0.810
MEAL	0.945	5.35	0.177
EA	5.80	7.48	0.775
CO	-56.50	7.42	-7.61
CONSTANT	126.00	22.00	5.75

NOTE: All values are authors' calculations. Number of observations = 323; R-squared = 0.871.

■ 4 Cited in Kahn (1988).

do not exhibit this property to a significant extent. For a flight of average length, first-class and coach fares increase about $0.22 per mile and discount fares increase about $0.06 per mile.

The PASS, SLOT, and HUB variables all measure possible capacity constraints facing the airlines serving a given route.[5] HUB is not statistically significant at the 5 percent level for any type of fares. The density of traffic on a route as measured by the PASS variable significantly increases discount fares. Only discount-fare passengers pay the expected premium for flying into slot-restricted airports. Flying into a slot-restricted airport increases the one-way fare by about $18 for these passengers.

TABLE 4

Discount Fare Estimates

Variable	Estimated Coefficient	Standard Error	T-Ratio
CARRIERS	-17.50	4.76	-3.67
CARRIERS^2	2.19	0.905	2.42
MILES	0.791E-1	0.961E-2	8.24
MILES^2	-0.140E-4	0.434E-5	-3.23
POP	-0.868E-3	0.829E-3	-1.05
INC	-0.411E-2	0.679E-3	-6.05
CORP	-1.06	0.203	-5.22
PASS	0.853	0.217E-3	3.93
STOP	-3.85	1.48	-2.60
SLOT	17.70	4.63	3.82
HUB	-3.50	2.16	-1.62
MEAL	1.80	2.21	0.813
EA	-10.60	3.04	-3.49
CO	-4.17	3.09	-1.35
CONSTANT	113.00	9.10	12.40

NOTE: All values are authors' calculations. Number of observations = 323; R-squared = 0.799.

Flight characteristics, such as the number of intermediate stops on the flight, influence coach and discount fares, but not first-class fares. Coach passengers pay about $7.60 for each stop, whereas discount-fare passengers actually get compensated about $3.85 for each stop. The fare charged does not seem to depend on whether the flight includes a meal.

The characteristics of the cities involved influence the fare charged to the various classes of passengers. The larger the population of the origin city, the lower the fare for all three classes of service, although this result is statistically significant at the 5 percent level only for first-class fares. The per-capita income variable seems to affect only discount fares significantly. Discount fares fall as incomes rise, indicating that higher-income passengers expect compensation in the form of lower fares for flying with discount tickets, other things being equal. The more important the city is as a business center (as measured by CORP), the higher the first-class and coach fares tend to be. Discount fares, on the other hand, are lower.

Continental charges significantly less than other carriers for first-class and coach service, other things being equal. Conversely, Eastern charges significantly less for discount service than other airlines, other things being equal.[6] Texas Air may own both of these carriers, but they appear to follow different criteria in setting fares. Keep in mind that these carrier-based fare differentials reflect differing cost and demand characteristics, including quality of service.

IV. Conclusion

An understanding of forces setting fares and the level of competition in the airline industry is crucial in order to formulate effective public policies for the industry. Some analysts have suggested that the ease of entry into most airline markets after deregulation increased the competitiveness of fares, even though the actual number of carriers is relatively small. We found that the number of airlines serving a route does influence the fares charged for all classes of service. Thus, the airline industry is not perfectly contestable even when very recent data are employed.

The benefits to passengers of adding an additional carrier on a typical route are still sizable, with fares declining until about four carriers are serving the route. This result is the strongest for discount fares. Fares on routes with four to five carriers are about $20 less than fares on routes with only one carrier, other things being equal. This is about a third of the average one-way discount fare.

■ **5** It is reasonable to consider whether both the number of carriers and the number of passengers on a route should be treated as endogenous variables in equation (1). Hausman specification tests were performed and indicate that in setting the fare on a given route, these variables can be treated as exogenous variables.

■ **6** We only report results that controlled for Continental and Eastern Airlines, because only these two carriers appeared to behave differently from the other carriers in setting fares.

Since deregulation, the airlines' clear goal has been to maximize their profits. Thus, they charge the highest fare possible on all their routes, with competition among existing carriers and the ease of entry of new carriers limiting how high those fares can be on a particular route. It is important that policymakers look at both the actual number of competitors and the ease of entry for a particular route. Since the number of carriers serving the typical route has risen since 1983—even if one allows for the recent merger wave—this suggests that the market for air fares remains fairly competitive, but that public policies to ease the entry of more carriers per route could lead to increased benefits for consumers.

In short, these findings suggest that the traditional concepts of market concentration, such as the number of competitors, are still relevant in assessing the amount of competition on a given route, even in the deregulated environment. Consequently, the antitrust laws that are applied to other industries are pertinent to the airline industry.

References

Bailey, Elizabeth E., Graham, David R., and Kaplan, Daniel P., *Deregulating the Airlines,* Cambridge, MA: The MIT Press, 1985.

Baumol, William J., Panzar, John C., and Willig, Robert D., *Contestable Markets and the Theory of Industry Structure,* New York: Harcourt Brace Jovanovich, Inc., 1982.

Borenstein, Severin, "Hubs and High Fares: Airport Dominance and Market Power in the U.S. Airline Industry," Institute of Public Policy Studies Discussion Paper No. 278, University of Michigan, March 1988.

Butler, Richard V. and Huston, John H., "Actual Competition, Potential Competition, and the Impact of Airline Mergers on Fares," paper presented at the Western Economic Association Meetings, Vancouver, British Columbia, July 1987.

Call, Gregory D. and Keeler, Theodore E., "Airline Deregulation, Fares, and Market Behavior: Some Empirical Evidence," in Andrew F. Daughety, ed., *Analytical Studies in Transport Economics,* New York: Cambridge University Press, 1985.

Congressional Budget Office, "Policies for the Deregulated Airline Industry," July 1988.

Hausman, J.A., "Specification Tests in Econometrics," *Econometrica,* November 1978, *46,* 1251–71.

Kahn, Alfred E., "I Would Do It Again," *Regulation,* 1988, *2.*

Morrison, Steven and Winston, Clifford, "Empirical Implications and Tests of the Contestability Hypothesis," *The Journal of Law and Economics,* April 1987, *30,* 53–66.

Official Airline Guides, *Official Airline Guide: North American Edition,* April 1, 1987, *13.*

Standard & Poor's Corporation, *Statistical Service: Security Price Index Record,* 1986 Edition.

U.S. Department of Commerce, Bureau of the Census, *State and Area Data Handbook,* 1984.

U.S. Department of Commerce, Regional Economic Measurement Division, "County and Metropolitan Area Personal Income, 1982-84," *Survey of Current Business,* April 1986.

U.S. Department of Transportation, Center for Transportation Information, *Origin and Destination City Pair Summary,* Data Bank 6, Computer Files.

Whalen, Gary, "Actual Competition, Potential Competition, and Bank Profitability in Rural Markets," *Economic Review,* Federal Reserve Bank of Cleveland, Quarter 3, 1988, 14–20.

Bank Capital Requirements and the Riskiness of Banks: A Review

by William P. Osterberg
and James B. Thomson

William P. Osterberg is an economist and James B. Thomson is an assistant vice president and economist at the Federal Reserve Bank of Cleveland.

Introduction

Banks are required to hold capital primarily as a buffer against future losses and in order to reduce the exposure of the deposit insurer. However, as regulators and researchers have recognized, changes in capital requirements also affect bank portfolio behavior. It is possible that increased capital requirements may lead banks to increase their riskiness and thus increase their expected losses or increase the potential exposure of the deposit insurer.

The object of this article is to show that the impact of increased capital requirements depends on the extent to which deposit costs reflect bank portfolio risk.[1] In particular, we show that with risk-based deposit insurance, the incentives to increase leverage or portfolio risk in response to an increase in bank capital requirements are reduced.

The article is organized as follows. First, we define bank capital and discuss the mechanisms through which it is intended to affect bank behavior. Next, we discuss the incentives for banks to decrease their capital buffer (increase their leverage). These incentives mainly stem from conflicts between the interests of creditors (depositors) and stockholders. We also show how these incentives are influenced by pricing deposit insurance. Previous research has shown that deposit insurance that is not adjusted for risk may encourage banks to increase their riskiness. We discuss previous research on the impact of increased capital requirements. We then present a model in which deposit costs are allowed to vary with risk, including the risk associated with leverage and, thus, with the capital buffer. By comparing our results with those of previous studies where explicit deposit costs do not vary with portfolio risk and leverage, we show that risk-based deposit insurance reduces the incentives to increase leverage or portfolio risk in response to an increase in bank capital requirements.[2] We also show that risk-based deposit

■ 1 For uninsured deposits, deposit costs are the interest rate banks have to pay on the deposits. For insured deposits, the cost of a dollar of deposits is the interest rate paid on the deposits, plus the per-dollar deposit insurance premium.

■ 2 Even though we do not assume correctly priced deposit guarantees, we do not get perverse effects from risk-based premiums (see Pyle [1983]) because we assume that the FDIC does not make relative pricing errors (that is, that it can measure risk and price it consistently).

insurance reduces the variance of earnings and the expected loss to the federal deposit guarantor when banks fail.

Functions and Definitions of Bank Capital

Regulators define bank capital in terms of book values. The regulatory definition of bank capital usually includes claims on bank profits (equity), reserves on loans or securities, and long-term subordinated debt. The primary function of bank capital is to serve as a cushion against unanticipated losses on assets, thereby ensuring the solvency of the bank. Bank capital is also used to finance asset purchases and thus bank growth. Minimum capital requirements (measured in terms of capital-to-asset ratios) constrain bank growth when it is costly to raise capital by issuing stock. Otherwise, if the rate of return on assets exceeds the cost of funds, banks would try to increase size as much as possible. In this article, we focus on how capital requirements affect bank risk, rather than bank size.

Incentives for Banks to Engage in Risky Behavior

While banks in some ways may be different from other firms, banks' incentives to engage in risky behavior are in some ways similar to the incentives of nonfinancial corporations. In particular, in the absence of conflicts between stockholders and bondholders (depositors), total bank value maximization and bank equity value maximization lead to identical results. However, as Jensen and Meckling (1976) argue, conflicts arise between stockholders and bondholders that cause total bank value maximization and equity value maximization to differ. By increasing the risk of the bank's portfolio or by increasing financial leverage, stockholders may be able to reduce the risk-adjusted value of the depositor's claim on the bank and thereby reallocate wealth from depositors to the stockholders. Wealth is reallocated if the reduction in the value of the bank is less than the reduction in the value of creditor claims on the bank. This type of conflict is referred to as an agency problem in the finance literature.

In most models of bank behavior, banks maximize the market value of equity and thus have the incentive to increase the portfolio variance. Because the value of equity cannot fall below zero but can increase without limit, stockholders will choose investments with a greater likelihood of high profits, regardless of the chance of loss. Unlike stockholders, bondholders receive only the promised amount if returns are high, but lose increasingly more as returns fall below the total amount of their claim. Thus, creditors have an incentive to control stockholder behavior.

Any analysis of the impact of capital requirements must also consider the banks' incentives to increase leverage (that is, to minimize their capital holdings). If the cost of raising funds from issuing stock exceeds the cost of raising funds from deposits, stockholders will prefer to increase their asset holdings via deposits and thus lower their capital ratios. Lower capital ratios (higher leverage) increase the probability of bankruptcy and thus of losses to creditors. The cost of raising funds from deposits is influenced by the pricing of deposit insurance. When deposit insurance is not priced so as to reflect bank risk, we refer to it as being "mispriced." We contend that it is the mispricing of deposit insurance that is at least partially responsible for an incentive for increased leverage. It is this incentive that makes capital requirements binding.

At least for nonfinancial corporations, it is common practice for bond covenants to contain restrictions on stockholder/manager behavior (see Smith and Warner [1979]). In fact, capital requirements and restrictions on bank portfolios can be viewed as bond covenants designed to protect the creditors. On the other hand, creditors may be protected if interest rates reflect bank risk. Risk- or leverage-related deposit rates could influence stockholders' incentives to increase portfolio risk or leverage.

It is an accepted conclusion that fixed-rate deposit insurance encourages risky behavior. Even if the deposit insurance agency adjusts the deposit insurance premium so that banks on average pay high enough premiums to cover expected losses, safe banks subsidize risky banks. In the absence of "correct" pricing of deposit insurance, and given the unresolved agency conflict between creditors and stockholders, banks will attempt to maximize the subsidy provided by the deposit insurance agency by increasing portfolio variance and leverage.[3] In this situation, there is a rationale for restrictions on bank leverage. However, if deposit costs reflect the increased risk associated with higher leverage, capital restrictions may no longer be necessary or binding.

■ 3 Correct pricing means that the deposit guarantor charges a deposit insurance premium equal to the risk premium the market would charge for uninsured deposits (see Thomson [1987]).

Mathematical Models of the Impacts of Increased Capital Requirements

Most mathematical models of the impacts of increased capital requirements assume that the bank is run for the benefit of the owners or stockholders. The creditors (depositors and deposit guarantors) are viewed as passive, perhaps being protected somewhat by bank portfolio restrictions designed to limit the ability of banks to engage in risky activities and the covariation of deposit costs with portfolio risk. Without an explicit model of either the creditors' position (for example, the market value of their claim) or the exact nature of the agency conflict, these analyses cannot explain the financial structure or capital position of the bank. The unresolved agency conflict pushes the capital-asset ratio towards its minimum.

The impact of capital regulation also depends on the overall regulatory structure. Both the difficulty of monitoring banks and uncertainty about the willingness of the guarantors to honor explicit and implicit guarantees play a role (see Kane [1986]). Pyle (1986) and Merton (1977) show how the value of deposit insurance depends on the closure policy and auditing frequency. Pennacchi (1987) shows how banks' preferences for greater leverage depend on the regulator's closure policy.

In our model, as well as in the models we survey below, the bank is closed at the end of a finite period of time. If the gross return on assets is insufficient to pay off depositors, the insurer provides the difference. In effect, these studies simplify the analysis by assuming that insolvent banks are closed and that there are no monitoring difficulties or uncertainties about closure.

A relatively early study by Koehn and Santomero (1980) viewed banks as utility-maximizers. They concluded that increased capital requirements would lead to increased asset risk, and possibly to increased risk of bank failure. However, interest rates did not reflect increased riskiness, as we would expect if deposits were uninsured. Neither was there an explicit treatment of deposit insurance. Keeley and Furlong (1987) emphasize the problems with the utility-maximization approach.

Karenken and Wallace (1978) utilize the state-preference framework and assume that the deposit rate is fixed. However, due to the presence of the deposit-insurance subsidy, the net deposit cost varies with asset risk and leverage. Lower leverage or lower asset risk decreases the probability of bankruptcy and hence the value of the subsidy.

A third approach utilizes the cash-flow version of the Capital Asset Pricing Model (Lam and Chen [1985]). Deposit interest rates vary but do not necessarily reflect asset risk or leverage. Hence, there may still be a subsidy provided by deposit insurance. Nonetheless, the covariation between deposit rates and the rate of return on assets plays a role in the bank's portfolio decisions. When deposit interest rates covary with the return on the bank's portfolio, the marginal return associated with increased asset risk or leverage is reduced. Therefore, the impact of increased capital requirements on bank risk and the probability of bankruptcy depends on whether interest rates are held fixed or whether they covary with the rates earned on assets.

Deposit Insurance Pricing

A separate body of research shows how deposit insurance should be priced. Merton (1977) models deposit insurance as a put option, showing how the value of the put option, and thus the position of the insurer, varies with the bank's leverage and portfolio risk.[4] Since increased leverage implies greater expected costs to the insurer, with correctly priced deposit insurance the premium charged each bank increases with bank leverage and portfolio risk, where the portfolio risk is measured as the variance of the earnings on assets. With correct pricing, there is no subsidy to the banks. Higher leverage results in higher insurance premiums, ameliorates the incentives to increase leverage, and modifies the impact of increased capital requirements.

I. The Joint Effects of Capital Requirements and Risk-Based Deposit Insurance on Optimal Bank Portfolios

The Model

In Osterberg and Thomson (1988) the cash-flow version of the Capital Asset Pricing Model (CAPM) used by Lam and Chen was modified to allow for an endogenously determined cost of deposits. The cost of deposits varies in a manner

■ 4 A put option is a contract that gives its holder the right to sell an asset at a predetermined price to the issuer of the option on or before a specified date. It represents a right but not an obligation to sell the asset.

similar to that suggested by the literature discussing the "correct pricing" of deposit insurance (for example, Merton [1977]). By comparing the results of our paper with those of previous studies where explicit deposit costs do not vary with portfolio risk and leverage (Lam and Chen [1985], and Koehn and Santomero [1980]), we show how risk-based deposit insurance changes the incentives to increase leverage or portfolio risk (as measured by the variance of earnings) in response to an increase in bank capital requirements.

The organization of the model and the basic results are presented below. As in our earlier paper, we make the usual assumptions necessary for the CAPM to hold. Furthermore, we assume that bankruptcy costs and taxes are zero and that the bank is operated by its owners.[5] The owners seek to maximize the value of bank equity, V, which has three components:

$$(1) \quad V = \frac{1}{R}\,[E(\tilde{\pi}) - \lambda CV(\tilde{\pi},\tilde{W}) - \lambda CV(\tilde{\pi},\tilde{\pi})], \text{ with}$$

$$CV(\tilde{\pi},\tilde{W}) = \sum_{j=1}^{n} A_j \sigma_{j,w},$$

$$CV(\tilde{\pi},\tilde{\pi}) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_i A_j \sigma_{i,j},$$

and A_i = amount invested asset i,
$\sigma_{i,j}$ = covariance between rates of return on asset i and j;
$\sigma_{j,w}$ = covariance between rates of return on asset j and cash flows of all other firms;
R = one plus the risk-free rate;
$\tilde{M}$ = aggregate cash flow of all firms in the market;
$\tilde{\pi}$ = cash profit of the bank;
$E(\tilde{\pi})$ = expected value of cash profit;
λ = market price of risk-bearing services;
$\tilde{W}$ = aggregate cash flow in the market, excluding the bank.

As in Lam and Chen (1985), the covariance between the cash profit of the bank and the aggregate cash flow of all firms, $CV(\tilde{\pi},\tilde{M})$ is partitioned into internal portfolio risk $CV(\tilde{\pi},\tilde{\pi})$ and external risk $CV(\tilde{\pi},\tilde{W})$ by separating the aggregate cash flows $\tilde{M}$ into $\tilde{\pi}$ and $\tilde{W}$, where $\tilde{W}$ is the aggregate cash flows in the market, excluding the bank. This allows us to isolate the risk of the asset portfolio (internal risk) from market risk in the maximization problem.

■ **5** The owner-manager assumption is used to resolve the agency problem that may exist between outside stockholders and managers (see Jensen and Meckling [1976]).

Suppose that there are N risky assets in which the bank can invest. Let $\tilde{r}_j$ be the uncertain return on asset j. Furthermore, the bank issues only insured deposits, D, and a fixed amount of capital, K. The bank pays its deposit guarantor (henceforth, the FDIC) a premium of g per dollar of deposits. Its expected cash profits at the end of the period are

$$(2) \quad E(\tilde{\pi}) = \sum_{j=1}^{n} \bar{r}_j A_j - (R+g)D.$$

The deposit insurance premium, g, varies with the bank's leverage and asset portfolio decisions (internal risk). We assume that the bank knows how its choices influence g, and thus what g results from its asset portfolio and capital structure decisions.

We can view the minimum ratio of deposits to capital, $C \equiv D/K$, as a covenant imposed on the bank by the FDIC in exchange for its deposit guarantees. A second restriction is the balance-sheet constraint that sources of funds must equal uses of funds. Thus, the problem facing the bank is to maximize V with respect to A_j and D, subject to

$$(3) \quad \sum_{j=1}^{n} A_j = D + K \text{ and}$$

$$(4) \quad D \leq CK \ (D = CK \text{ when the capital constraint is binding}).$$

The solution to this problem is a series of optimality conditions describing the bank's choices (see Osterberg and Thomson [1988]). We assume that the capital constraint is binding and thus that equity value could be increased with a looser capital requirement. The bank will choose its asset mix so that marginal expected returns of all assets are equal. The marginal increase in equity value from a lower capital requirement, γ, is just equal to the risk-adjusted return on assets less the cost of deposits. Changes in leverage and portfolio composition affect γ.

We utilize Merton's (1977) put option formulation of FDIC deposit insurance, which indicates how g varies with portfolio variance (ρ) and leverage (δ). ρ and δ are nonnegative functions of portfolio variance and leverage, respectively. We do not assume, however, that the deposit guarantor correctly prices the insurance and drives the net value of the FDIC's claim to zero (see Osterberg and Thomson [1987]). As a result, the agency problem is not completely resolved, and the stockholders still have incentives to increase leverage and portfolio risk (hence the binding capital constraint).

Bank stockholders seek to maximize equation (1) subject to (3) (the balance-sheet constraint) and (4) (the capital constraint). The optimality conditions, from the constrained maximization problem, for the n assets can be written as (see Lam and Chen [1985] or Osterberg and Thomson [1988])

$$(5)\quad 2[\lambda + \rho CK]\sum_{i=1}^{n} A_i\sigma_{i,k} + R\gamma + CK\delta = \alpha_k - R - g, \qquad (k = 1,2,\ldots,n).$$

The right side of (5) represents the expected spread associated with investing in asset k. α_k is the return on asset k adjusted for external risk. γ is the Lagrangian multiplier associated with a binding capital constraint. Note that the risk-based deposit insurance premium affects portfolio decisions by affecting the spread of return over cost and by affecting the risk adjustment associated with changes in leverage and variance.

Portfolio Composition

As in Osterberg and Thomson (1988), the solutions for the multiplier, γ, and the optimal portfolio shares, A_k^*, are

$$(6)\quad \gamma = \{[2(\lambda + CK\rho)]^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}\}^{-1} \{[2(\lambda + CK\rho)]^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}[\alpha_i - R - g - CK\delta] - (1 + C)K\}.$$

$$(7)\quad A_k^* = [2(\lambda + \rho CK)]^{-1}\{\sum_{j=1}^{n} v_{k,j}\alpha_k - R - g - CK\delta - \gamma\} \quad (k = 1,2,\ldots,n).$$

Here $v_{i,j}$ is the ijth element of the inverse variance-covariance matrix of the asset shares A_j.

Let γ_F and A_k^F be the multiplier and the optimal asset share for the fixed-rate deposit insurance case (that is, $g = \overline{g}$, $\rho = 0$, and $\delta = 0$). Equations (6) and (7) can be rewritten as

$$(6a)\quad \gamma = \gamma_F - CK\delta - CK\rho(1 + C)K,$$

$$(7a)\quad A_k^* = \frac{2\lambda A_k^F + \rho CK(1 + C)K}{2(\lambda + \rho CK)}$$

Note that γ is smaller under risk-based deposit insurance than under fixed-rate deposit insurance because by definition C, K, δ, and ρ are positive.[6] γ can be interpreted as the cost to the bank of a more restrictive capital constraint. In this model, the γ is positive because of agency problems. By tying deposit costs to bank-asset risk and leverage, the risk-based deposit-insurance premiums in this model partially resolve the agency conflict and, hence, lower the cost of the capital constraint.[7] Intuitively, deposit rates that do not vary with risk or leverage provide a subsidy to the stockholders. The subsidy increases with the risk and leverage of the bank. Risk-based deposit rates reduce the risk- and leverage-related subsidy and therefore the cost to stockholders of increasing the capital constraint.

Equation (7) shows that the optimal portfolio share for asset k is a function of γ. Since γ is smaller for banks paying risk-based deposit rates than for banks paying fixed-rate deposit rates, the impact of the capital requirements has less impact on portfolio composition for banks paying risk-based premiums than for banks paying fixed-rate premiums. Equation (7a) gives the relationship between the optimal portfolio share for asset k under fixed- and variable-rate premiums. From (7a) it is clear that adjusting deposit-insurance premiums for asset risk and leverage has an uncertain impact on portfolio composition. To see more clearly the effects of risk-based premiums on portfolio composition, we substitute (6) into (7),

$$(7b)\quad A_k^* = [2(\lambda + \rho CK)]^{-1}\{\sum_{j=1}^{n} v_{k,j}\alpha_k - \frac{\sum_{j=1}^{n} v_{k,j}}{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}\alpha_k\} + \frac{(1 + C)K\sum_{j=1}^{n} v_{k,j}}{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}} \quad (k = 1,2,\ldots,n).$$

If we set ρ equal to zero in (7b) we get A_k^* for a bank paying fixed-rate deposit-insurance premiums.

■ **6** This differs from Lam and Chen's stochastic interest-rate case where the capital constraint multiplier may be larger or smaller than the capital constraint multiplier in the deterministic deposit case.

■ **7** The risk-based deposit-insurance premiums only partially resolve the agency conflict because we do not assume the FDIC charges the bank the full value of the insurance. That is, we do not impose correct pricing on the model.

From (7b) the optimal asset share is a function of the expected asset returns adjusted for outside risk weighted by the elements of the inverse of the variance-covariance matrix. The fixed-rate deposit insurance result is identical to Lam and Chen's result when Regulation Q prevails and is equivalent to Koehn and Santomero's results. For both fixed-rate and risk-based deposit insurance, A_k^* is also a function of the capital constraint. When variable-rate deposit insurance is introduced into the model, A_k^* is also a function of the change in the cost of deposit insurance due to a change in the risk of the bank's portfolio, ρ. It is interesting to note that A_k^* is not a function of δ or g.

The impact of increased capital requirements on asset portfolio composition is uncertain for banks facing both the fixed-rate and risk-based deposit insurance. The indeterminate sign on $\frac{\partial A_k^*}{\partial C}$ is consistent with the findings of Lam and Chen.[8] That is, although the purpose of an increase in the capital requirement is to reduce overall bank risk, it may cause the bank to choose a riskier portfolio and may increase overall bank risk.

Portfolio Risk and Expected Profits

For investors and bank regulators, it is not the risk or return of the individual activities (or assets) that matters, it is the risk-adjusted return on the bank's portfolio. Therefore we are interested in the effects of risk-based deposit insurance and changes in capital requirements on internal risk (portfolio risk), $CV(\tilde{\pi}, \tilde{\pi})$, and on expected profits, $E(\tilde{\pi})$. From Osterberg and Thomson (1988), the portfolio risk and the expected profits of the optimal bank portfolio are

$$(8)\quad CV(\tilde{\pi},\tilde{\pi}) = (2[\lambda + \rho CK])^{-2}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}\alpha_i\alpha_j - \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}\alpha_i}{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}\alpha_j\right\} + \frac{[(1+C)K]^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}}$$

$$(9)\quad E(\tilde{\pi}) = (2[\lambda + \rho CK])^{-1}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}r_i\alpha_j - \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}r_i}{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}\alpha_j\right\} + \frac{(1+C)K\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}r_i}{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{i,j}} - (R+g)CK.$$

If we set $\rho = 0$, equation (8) is the variance of earnings in the fixed-rate deposit case. Note that like A_k^*, $CV(\tilde{\pi}, \tilde{\pi})$ is not a function of δ or g. Furthermore, because ρ is positive, the variance of portfolio earnings for a bank with fixed-rate deposit insurance is greater than the variance of earnings for a bank with risk-based deposit insurance. In other words, banks that have to pay depositors (or the FDIC) for risk-bearing services will hold less-risky portfolios than banks that do not have to pay for those risk-bearing services. This result holds for all values of C.

As in Lam and Chen, an increase in the capital requirement leads to a reduction in portfolio risk under fixed-rate deposit insurance. That is, $\frac{\partial CV(\tilde{\pi},\tilde{\pi})}{\partial C}$ is positive when $\rho = 0$. However, the sign of $\frac{\partial CV(\tilde{\pi},\tilde{\pi})}{\partial C}$ is ambiguous for banks facing risk-based premiums. Therefore, the joint effect of a more restrictive capital constraint and of risk-based insurance premiums may be to increase bank portfolio risk.[9] However, because the value of (8) is greater when banks face fixed-rate premiums than when they face risk-based premiums for all C, risk-based premiums result in less internal risk than do fixed-rate premiums. Therefore, so long as the FDIC does not make relative errors in pricing its guarantees, risk-based deposit-insurance premiums do not introduce any new perverse effects into the analysis.

If we set $g = \overline{g}$ and $\rho = 0$, equation (9) is the expected profits for a bank with fixed-rate deposit insurance. As anticipated, when the risk

■ **8** Lam and Chen also get an indeterminate result for the net effect of more stringent capital requirements on overall bank risk in their stochastic deposit case.

■ **9** Separation between capital structure and portfolio decisions may not hold in our model because we do not assume that the deposit guarantor charges banks a premium equal to the fair value of the deposit guarantees.

profile of the bank results in a risk-based premium, g, equal to the fixed-rate premium, $\overline{g}$, profits are lower for the bank paying risk-based premiums than for the bank paying fixed-rate premiums. This result holds because, as we know from equation (8), banks paying fixed-rate premiums will hold riskier portfolios than banks paying risk-based premiums, and there is a positive relationship between risk and return (expected profits).

For both fixed-rate and risk-based insurance, the effect of a change in C on expected profits is ambiguous. Since expected profits are not adjusted for risk, it is possible for a relaxation of the capital constraint to increase the value of the firm and to reduce profits. This result was also found by Lam and Chen (1985).

Bankruptcy Risk

The only time the FDIC must honor its guarantees is when a bank fails. So, the impact of changing the capital requirement on the risk of bankruptcy is an important issue for the FDIC. A bank's bankruptcy risk is a function of asset portfolio risk and leverage. Since an increase in the capital requirement reduces leverage, an increase in internal risk in response to increased capital requirements does not necessarily increase bankruptcy risk. Koehn and Santomero (1980) show that the probability of failure, P, is

$$(10) \quad P = Pr\{\tilde{\pi} < K\} \leq \frac{CV(\tilde{\pi},\tilde{\pi})}{[E(\tilde{\pi}) - K]^2}$$

Holding C constant, the impact of risk-based deposit insurance is to reduce both the numerator and denominator of P. Therefore, the impact of risk-based insurance on default risk is uncertain. On the other hand, a reduction in the variance of earnings should reduce the expected loss to the FDIC when a bank fails. From this standpoint, risk-based deposit insurance produces a desirable result.

Lam and Chen (1985) show that the impact of changing the capital requirement on P is indeterminate for fixed-rate deposit insurance. It is also indeterminate when risk-based deposit insurance is introduced. Our inability to sign $\frac{\partial P}{\partial C}$ for banks with risk-based deposit insurance is at least partially due to our assumption that the FDIC does not charge banks for the fair value of their insurance.

II. Conclusion

Studies of the impact of changes in capital requirements on bank portfolio behavior and risk are extremely sensitive to the assumptions of how deposit insurance is priced. Previous mathematical analyses of the impact of increased capital requirements on bank portfolio behavior implicitly or explicitly assume that deposit insurance is mispriced. This introduces an agency problem into the analysis that causes the capital constraint to be binding and generates the conclusions of these studies. We contend that with correct pricing of deposit insurance the capital constraint is no longer binding. Using a modified version of the cash-flow CAPM, which incorporates a put option formulation for deposit insurance, we compare the results of our earlier study (Osterberg and Thomson [1988]), where deposit rates vary with portfolio risk and leverage, to the general results of previous studies where explicit deposit costs are independent of portfolio risk and leverage.

We find that, with risk- and leverage-related deposit rates, the incentive to increase leverage is smaller than when the deposit rate and insurance premium are fixed. Allowing explicit deposit costs to vary with risk and leverage also reduces the portfolio variance. In addition, asset choice is influenced by the response of the risk premium to increases in portfolio variance.

As in the case where explicit deposit costs do not vary with risk and leverage, the impact of increased capital requirements on portfolio behavior for banks paying risk-based deposit insurance premiums is generally ambiguous. In both cases, the impact of increased capital requirements on asset choice is indeterminate, as are the responses of portfolio variance, expected profits, and the probability of bankruptcy. However, our failure to impose correct pricing may be responsible for these indeterminacies. Nonetheless, allowing deposit rates to vary with portfolio risk and leverage results in reductions in portfolio variance and in the incentive to increase leverage. These would seem to be desirable results from a regulator's viewpoint.

References

Jensen, Michael C. and Meckling, William H., "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics,* October 1976, *4,* 305–60.

Kane, Edward J., "Appearance and Reality in Deposit Insurance: The Case for Reform," *Journal of Banking and Finance,* June 1986, *10,* 175–88.

Karenken, John and Wallace, Neil, "Deposit Insurance and Bank Regulation: A Partial Equilibrium Exposition," *Journal of Business,* July 1978, *51,* 413-38.

Keeley, Michael C. and Furlong, Frederick T., "Bank Capital Regulation: A Reconciliation of Two Viewpoints," *Working Paper* No. 87-06, Federal Reserve Bank of San Francisco, September 1987.

Koehn, Michael and Santomero, Anthony, "Regulation of Bank Capital and Portfolio Risk," *Journal of Finance,* December 1980, *35,* 1235–44.

Lam, Chun H. and Chen, Andrew H., "Joint Effects of Interest Rate Deregulation and Capital Requirements on Optimal Bank Portfolio Adjustments," *Journal of Finance,* June 1985, *40,* 563–75.

Merton, Robert C., "An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees," *Journal of Banking and Finance,* June 1977, *1,* 3–11.

Osterberg, William P. and Thomson, James B., "Deposit Insurance and the Cost of Capital," *Working Paper* 8714, Federal Reserve Bank of Cleveland, December 1987.

———, "Capital Requirements and Optimal Bank Portfolios: A Reexamination," *Proceedings: A Conference on Bank Structure and Competition,* Federal Reserve Bank of Chicago, 1988.

Pennacchi, George G., "A Reexamination of the Over- (or Under-) Pricing of Deposit Insurance," *Journal of Money, Credit and Banking,* August 1987, *19,* 340–60.

Pyle, David H., "Pricing Deposit Insurance: The Effects of Mismeasurement," *Working Paper* 8305, Federal Reserve Bank of San Francisco, October 1983.

———, "Capital Regulation and Deposit Insurance," *Journal of Banking and Finance,* June 1986, *10,* 189–201.

Smith, Clifford W. and Warner, Jerold B., "On Financial Contracting: An Analysis of Bond Covenants," *Journal of Financial Economics,* June 1979, *7,* 117–61.

Thomson, James B., "The Use of Market Information in Pricing Deposit Insurance," *Journal of Money, Credit and Banking,* November 1987, *19,* 528–37.

Turnover, Wages, and Adverse Selection

by Charles T. Carlstrom

Charles T. Carlstrom is an economist at the Federal Reserve Bank of Cleveland. This paper is based on the author's Ph.D. dissertation at the University of Rochester. The author wishes to thank Randall Eberts, William Gavin, Erica Groshen, James Hoehn, Kenneth McLaughlin, Walter Oi, and Richard Rogerson, all of whom provided helpful comments on earlier drafts of the paper.

Introduction

Worker mobility is necessary for the efficient operation of the labor market, so that the best matches can be found between workers and employers. Employers have only limited information about the abilities of each prospective worker, however. When making hiring decisions, they take the chance of employing a worker who does not have the skills (and thus the productivity) that was originally expected.

Both high- and low-productivity workers seek higher-paying jobs at any given time. The problem facing the employer is how to distinguish between the two. Low-productivity job searchers, of course, try to pass themselves off as high-productivity workers. The employer can discern a worker's true abilities only after the hiring decision has been made, however. Because of this asymmetrical information, workers' ability to change jobs and find the best match may be seriously impaired. Consequently, the labor market may not work efficiently.

This paper suggests that asymmetrical information can result in adverse selection. Adverse selection is a term coined by Akerlof (1970) to explain why the used-car market is dominated by "lemons." Car owners, he argues, often sell their vehicles because of poor performance or unreliability. Potential buyers realize the owner's motivation and pay less for a used car because of the likelihood of purchasing a lemon. The tendency is then reinforced for new-car owners to sell their vehicles only if they are unreliable. Thus, adverse selection can explain why a new automobile sells for considerably less as soon as it is driven off the showroom floor. In the case of the labor market, adverse selection comes about because low-productivity workers may change jobs in order to be confused with high-productivity job-changers.

A model of worker mobility based on adverse selection can help to explain several stylized facts of the labor market, particularly in regard to job turnover and wages. First, as Mincer (1984) shows, frequent job mobility among older workers results in lower wages. Second, while earnings for all workers tend to increase over time, older workers who quit generally experience zero or negative wage growth (Bartel and Borjas [1981]). Adverse selection can also help to explain why workers who have had a history of frequent job moves are more likely to move in the future.

For the same reason that lemons may dominate the used-car market, lower-productivity

workers tend to be frequent job-changers. These workers will then have lower wages, on average, compared to infrequent job-changers. This can explain why mobility among older workers results in lower wages and why prior mobility can predict future mobility.

These empirical regularities are frequently explained by combining the concept of "firm-specific" human capital with the assumption that workers differ in their propensities to change jobs. Firm-specific human capital is knowledge that increases workers' productivity at their present firm, but that cannot be transferred to other firms. Thus, as a worker's tenure with the same firm increases, his or her firm-specific knowledge grows, pushing up his or her productivity.

Frequent job-movers would invest in less firm-specific human capital, since the knowledge they gain on the job would be forfeited after each job change. The argument is then that frequent job-movers would have lower average wages and flatter age-earnings profiles than infrequent job-changers. Consequently, infrequent job-movers would have a steeper age-earnings profile than would the frequent job-movers.

Arguments based on firm-specific human capital have some problems explaining these observations of the labor market, however. First, no reasons are given for the assumed difference in propensities among workers to change jobs. Second, the firm-specific human capital model cannot explain the relationship between wages and turnover in light of work by Salop and Salop (1976). They show that if a worker's propensity to move is not public information, then the infrequent job-changers would post bonds at firms in order to separate themselves from the frequent job-movers. The implication is that the wage rates for job-movers should be higher than those of job-stayers in the early part of their careers—an observation that is inconsistent with the findings of Bartel and Borjas.

Third, if a substantial number of firms require general rather than firm-specific human capital, then frequent movers would sort themselves into these firms. They thus would have steeper age-earnings profiles because they would bear the full cost of acquiring the general human capital. In firms with primarily firm-specific human capital, the costs would be shared by both the worker and the employer.

The adverse-selection model of labor mobility can help explain these empirical anomalies without relying on firm-specific human capital. Moreover, it provides a basis for examining the welfare implications of "lemons" (low-productivity workers) in the labor market. The model predicts that mobility is hampered because frequent moves can brand a person as a low-productivity worker. However, it is shown that government subsidies to increase mobility would be ineffective.

The model is developed in several steps. The basic assumptions are presented in section I. The first version of the model incorporates contingent wage contracting, which in effect allows firms to sort among workers according to productivity. In this case, it is shown that adverse selection is not a problem, since all workers are paid their expected output.

The remaining versions of the model exclude contingent wage contracts, which introduces the situation in which all workers receive the same wage, ex ante. This pooling of low- and high-productivity workers creates adverse selection, where the low-productivity workers are the frequent job-changers.

Example 2 of the model assumes that workers post no bonds, in which case the worker's wage in every period is the firm's estimate of his or her productivity. Next, example 3 allows bonding, which benefits high-productivity workers (the infrequent job-changers) and hurts low-productivity workers (the frequent job-changers). Bonds arise in order for firms to compete for the high-productivity workers. Finally, example 4 is a two-period model with bonding. This example is useful for discussing the welfare implications of the model, which are presented in section II.

I. Job Mobility and Adverse Selection

In the following model of the labor market, workers are assumed to live and work for three periods indexed by 1, 2, and 3. A three-period horizon allows the model to explain why workers who moved frequently in the past are more likely to move in the future. At the end of periods one and two, workers decide whether to continue working at their present job or to change jobs. Workers change jobs if the change raises the expected value of their future wages.

Productivity at a firm consists of a job-matching component, θ, and an individual-specific component, p. The labor contribution to production is represented by the simple linear relation $y = p + \theta$. A worker's base productivity level, p, is assumed to be constant across firms. The matching component, θ, varies across firms, so workers shop around in order to improve their job match. However, since only the workers know their own base productivity, p, firms cannot immediately observe whether a new worker changed jobs because he was a high-productivity worker with a bad job match or because he was

a low-productivity worker wishing to be confused with a high-productivity job-changer.

The following restrictions are placed on the distributions of p and θ: p is assumed to be distributed on the interval $[p', p'']$ with a cumulative distribution function of $F(p)$ and a density function of $f(p)$; θ is assumed to be distributed on the interval $[-\theta', \theta'']$ with a cumulative distribution function of $G(\theta)$ and a density function of $g(\theta)$ with $E(\theta) = 0$. In addition, it is assumed that θ is independent both across individuals and across different jobs, and that p and θ are independently distributed random variables. Thus, a worker's current job match—or the quality of another worker's match—does not provide the worker with any information regarding his match at another firm. Similarly, a worker's productivity does not indicate which job or task he will be most productive in performing.

Prior to production, neither firms nor workers know what θ will be, although workers know their own productivity type, p. After one period, a worker's output at the firm, y, is assumed to be perfectly observable by both the worker and the firm. Furthermore, it is assumed that a worker's output at a firm is constant over time but cannot be observed by other firms.[1]

For simplicity, it is assumed that firms cannot observe an applicant's past wage rates. This ensures that workers who did not move after the initial period will not move in subsequent periods. The only reason a worker would want to move after the second period would be to find a better job match. He would not move after period two, however, because the incentive to search for a better match declines with age.

Example 1: Mobility and Wages With Contingent Wage Contracts

This section examines the model's properties in an economy with no restrictions on types of wage contracts offered, in order to show that the stylized facts of the labor market cannot be explained without adverse selection. The model predicts that workers will be paid their realized output, y, at the end of each period. This is called a contingent wage contract, because a worker's pay is contingent on his or her realized output in that period.

■ 1 This assumption is not crucial because observing a worker's output at a previous firm would give a potential employer a "noisy" signal of a worker's base productivity level.

Since workers are risk-neutral, they are indifferent between accepting a wage equal to their base productivity level, p, or accepting a wage equal to their realized output, y. Contingent wage contracts in effect allow firms to sort among workers according to productivity. If workers are paid based on their output, the model collapses to a simple version of a standard job-matching model, in which workers move only to seek better matches.

Define $W_1(p)$ to be the value of future wage payments at the beginning of the first period for a worker with a base productivity level of p; define $V_2(y)$ to be the value of future wage payments at the beginning of period two for a worker who produced $y = p + \theta_1$ in the first period and decided not to move; and define $\lambda_2(p, \theta_2)$ to be the value of future wage payments for a worker who moved after the first period.

$$(1) \quad W_1(p) = p + \theta_1 + E_1 \max[\lambda_2(p, \theta_2), V_2(y)]$$

where

$$\lambda_2(p, \theta_2) = p + \theta_2 + E_2 \max[p + \theta_3, p + \theta_2]$$

$$V_2(y) = 2(p + \theta_1)$$

and

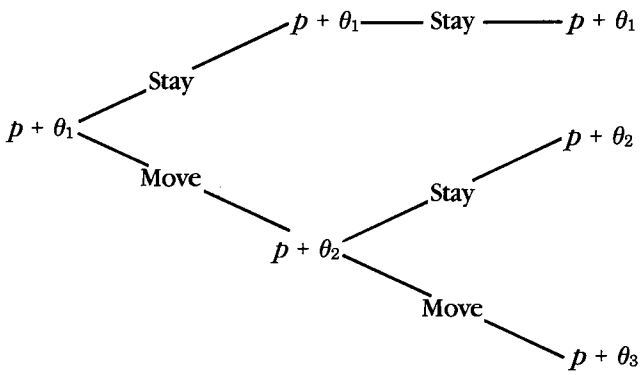
θ_i = match at ith firm,
E_t = expectation given the information at the end of period t.

$W_1(p)$ consists of the worker's first-period wage (the value of his productivity $p + \theta_1$) and either λ_2 or V_2, depending on whether he switches jobs after the first period. A worker switches jobs if $\lambda_2 > V_2$, but stays at his job if $\lambda_2 \leq V_2$. If a worker does not move after period one, he earns his output, $y = p + \theta_1$, in both periods two and three. A worker who moves after period one will earn $p + \theta_2$ in the second period and then either his output, $p + \theta_2$, if he stays and works at this firm again in period three, or p if he switches jobs once more. A worker who changes jobs after the first period will do so again if his output, $p + \theta_2$, is less than p (his expected wage if he moves). Thus, a worker who changes jobs after the first period does so again if $\theta < 0$. Figure 1 depicts a worker's wage based on whether he moves or stays at his firm after periods one and two.

The reservation output level for a p-productivity worker, $y^r(p)$, is defined to be the wage at which the worker is indifferent between staying and leaving, $V_2(y^r(p)) = E_1 \lambda_2(p, \theta_2) = \lambda_2(p)$. A worker stays at his present job if $p + \theta_1 > y^r(p)$. This definition implies the following

FIGURE 1

Possible Moves and Wages With Contingent Contracts



SOURCE: Author.

expressions for the expected values of W, V, and λ (where E denotes the expectations operator):

$$(2) \quad EW_1(p) = p + G(y^r(p)-p)\lambda_2(p) + (1-G(y^r(p)-p)) \times E(V_2(p+\theta_2 | \theta_2 > y^r(p)-p)$$

$$(3) \quad E\lambda_2(p) = p + G(0)p + (1-G(0))(p + E(\theta_2 | \theta \geq 0))$$

$$(4) \quad EV_2(y) = 2p.$$

The probability that a worker leaves after the first period, $\theta_1 < y^r(p) - p$, is given by $G(y^r(p)-p)$, since G is the cumulative distribution function of θ. Similarly, $G(0)$ is the probability that $\theta_2 < 0$ and is the probability that a worker will move after period two if he moved after the first period. The expected value of future wages for a worker who moved after the first period, $\lambda_2(p)$, is his expected wage in the second period, p, plus the product of his probability of moving again, $G(0)$, and his average wage if he moves again, p, plus the probability that he does not move, $(1-G(0))$, multiplied by his expected wage if he stays, $p + E(\theta_2 | \theta_2 \geq 0)$.

If we further assume that the θs are uniformly distributed over the interval $(-\theta', \theta')$, then the reservation output for a risk-neutral p-productivity worker with no search costs is $y^r(p) = p + \theta'/8$. The probability that a worker moves after period zero would then be $G(y^r(p)-p) = 9/16$, and the probability that a worker moves after period two, given he moved after period one, would be $G(0) = 1/2$. These separation probabilities are constant across workers, implying that adverse selection is not a problem. The reason is that, on average, workers are paid their expected output, p.

The example predicts that job-movers—those with the worst matches—earn lower wages. However, it cannot explain why these same workers have less future wage growth. Similarly, the driving force behind this result is the matching characteristic, which can explain the mobility of younger workers. However, it cannot explain the empirical evidence which suggests that older workers, but not younger workers, are hurt when moving.

Because most wage contracts are not contingent on a worker's future output, the remaining examples in this paper exclude contingent wage contracts. This introduces a pooling equilibrium, where, ex ante, all workers receive the same wage. The result is adverse selection, where the low-productivity workers are the frequent job-changers.

Adverse selection can explain why older workers are seemingly worse off after they move. Although a job-matching model is not realistic when considering the mobility of older workers, the assumption is maintained in order to ensure that some workers always change jobs. The matching component is not necessary for the following examples.[2] The next example examines the implications of the model excluding both contingent wage contracts and bonds, so that a worker's wage in every period is his expected output in that period.

Example 2: Mobility and Wages Without Contingent Wage Contracts and Bonds

The examples given in tables 1–4 assume that there are two types of workers, who can have three possible outputs at a firm. Half of the workers are high-productivity with $p = 2$; the rest are low-productivity with $p = 1$. The job-matching component is assumed to take on three values (–1, 0, or 1), each of which occurs with a one-third probability.

This example considers an equilibrium where no bonds are posted, that is, where a worker's wage in every period is the firm's estimate of his or her productivity. This implies that there will be a pooling equilibrium and that all workers will receive the same wage in the first period. With these assumptions, the solution given in tables 1 and 2 can be verified.

■ 2 See Greenwald (1986).

TABLE 1

Mobility and Wages for a Low-Productivity Worker Without Bonding

	Period 1	Period 2	Period 3
Output at which a low-productivity worker is indifferent between moving and staying	$y^r(1) = 1.4$	$w_3 = 6/5$	—
Fraction of workers who move at end of period	2/3	2/3	—
Wages for a low-productivity worker who never moves	3/2	$y = 2$	$y = 2$
Wages for a low-productivity worker who moves only after period one	3/2	$w_2 = 4/3$	$y = 2$
Wages for a low-productivity worker who moves after both periods one and two	3/2	$w_2 = 4/3$	$w_3 = 6/5$

NOTE: w_2 = the second-period wage for workers who changed jobs after period one; w_3 = the third-period wage for workers who changed jobs after periods one and two.
SOURCE: Author.

The transition probabilities and wages given in tables 1 and 2 can be shown to solve the preceding problem. First assume that the separation rates in the tables are correct. They can then be used to verify the wages, w_2 and w_3. Given that the wages are consistent with the separation rates, it is then necessary to show that these wages imply the separation rates posited.

For example, if the reservation output for a high-productivity worker is 1.7, then he will leave his original firm if $y < 1.7$ or equivalently if $y_1 = 1$, which occurs one-third of the time. High-productivity workers who stay will earn their output, which is either $y = 2$ or $y = 3$. If a high-productivity worker moved after period one, he would move again if $y_2 < w_3 = 6/5$. This occurs one-third of the time, or when $y_2 = 1$.

Similarly, low-productivity workers will move two-thirds of the time given their reservation outputs. With these transition probabilities, we can calculate the wages of job-movers. Then, λ_2 and $V(y)$ can be calculated with these wages to verify that the reservation output for a low-productivity worker, $y^r(1)$, is 1.4, while the reservation output for a high-productivity worker, $y^r(2)$, is 1.7.

Notice that the low-productivity workers move twice as often as the high-productivity workers: two-thirds (one-third) of the low- (high-) productivity workers move after period one, while two-thirds (one-third) of those who moved previously move again after period two. This is a result of adverse selection.

Example 3: Mobility and Wages Without Contingent Contracts, Bonding Allowed

Because of the difference in mobility between high- and low-productivity workers, example 2 cannot be an equilibrium once bonding is allowed. Firms could earn positive profits by trying to compete for the high-productivity workers, since firms make money by employing these workers and lose money by employing low-productivity workers.

Because high-productivity workers move only half as often as low-productivity workers, firms try to attract the high-productivity workers by requiring incoming workers to post bonds that are paid according to their future mobility. Those who change jobs forfeit their bonds, while the job-stayers split the proceeds of the bonds.

Bonding implies that workers no longer earn their expected productivity every period: instead, they are paid less than their expected productivity in the first period of an employment contract, and make up for this loss in later periods. The amount of the bond is the difference between a worker's expected productivity and his wage during the first period of an employment contract. In later periods, a worker is paid more than his marginal productivity, the bonus being the difference between his wage and his expected productivity.

Bonding benefits the high-productivity workers—those who move infrequently—and hurts the low-productivity workers—the frequent job-movers. Because bonds offset some of the income gained by the low-productivity workers as a result of adverse selection, they redistribute income from the low-productivity workers to the high-productivity workers. Competition for high-productivity workers ensures that workers post bonds, although in equilibrium, bonding may not be sufficient to separate workers according to their respective productivities.

TABLE 2

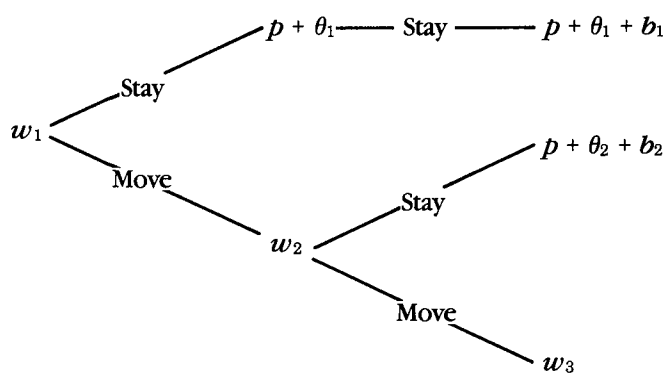
Mobility and Wages for a High-Productivity Worker Without Bonding

	Period 1	Period 2	Period 3
Output at which a high-productivity worker is indifferent between moving and staying	$y^r(1) = 1.7$	$w_3 = 6/5$	—
Fraction of workers who move at end of period	1/3	1/3	—
Wages for a high-productivity worker who never moves	3/2	$y = 2$ or 3	$y = 2$ or 3
Wages for a high-productivity worker who moves only after period one	3/2	$w_2 = 4/3$	$y = 2$ or 3
Wages for a high-productivity worker who moves after both periods one and two	3/2	$w_2 = 4/3$	$w_3 = 6/5$

NOTE: w_2 = the second-period wage for workers who changed jobs after period one; w_3 = the third-period wage for workers who changed jobs after periods one and two.
SOURCE: Author.

FIGURE 2

Possible Moves and Wages Without Contingent Contracts



SOURCE: Author.

Define b_1 to be the bonus paid to workers who did not change jobs after period one, and define b_2 to be the bonus paid to workers who switched jobs after period one and stayed after period two. Figure 2 depicts a worker's wage based on whether he moves or stays at his firm after periods one and two. Given the structure of bonding as described above, tables 3 and 4 illustrate the solution for this example.[3]

Tables 3 and 4 are an equilibrium for this example, since a potential firm could never successfully compete for either a low-productivity or a high-productivity worker. The low-productivity workers are still being confused with the high-productivity workers and thus do better than they would if they admitted that they were low-productivity workers and were paid their expected output, 1, every time they moved and did not post any bonds.

It can also be shown that if the amount of the bond posted by workers changed, the high-productivity workers would be made worse off.[4] This is because the bonuses, b_1 and b_2, are the largest possible so that the high-productivity workers still move. (That is, b_1 and b_2 are chosen such that a high-productivity worker who produces an output of 1 would be indifferent between moving and staying.) If b_1 were increased, high-productivity workers would never move, even if they have a bad match, $\theta = -1$. If b_2 were increased, high-productivity workers would never move after period two and would be made worse off.

This example illustrates that adverse selection is present in the model, since two-thirds (one-third) of the low- (high-) productivity workers

■ **3** Since two-thirds (one-third) of the (low-) high-productivity workers move after period zero and again after period one, the expected productivity of a worker who changes jobs after the first period is [(2/3 x 1/2 x 1) + (1/3 x 1/2 x 2)] / [(2/3 x 1/2) + (1/3 x 1/2)] = 4/3; the expected productivity of a worker who changes jobs after both periods is [(2/3 x 2/3 x 1/2 x 1) + (1/3 x 1/3 x 1/2 x 2)] / [(2/3 x 2/3 x 1/2) + (1/3 x 1/3 x 1/2)] = 6/5; and the expected productivity of a worker in the initial period is simply [(1/2 x 1) + (1/2 x 2)] = 3/2. The wages reported in the text can be obtained as follows. In the first period, the probability that a worker stays at his present job is 1/2, therefore $w_1 = 3/2 - (1/2)_1 b_2 = 10/9$; similarly, the conditional probability that a worker changes jobs after the second period given that he changed jobs after the first period is 4/9, therefore $w_2 = 4/3 - (4/9)\ {}_0b_2 = 56/45$; and the wage for a worker who changes jobs twice is his expected productivity, $w_3 = 7/6$.

■ **4** Under the assumptions of this model, bonds cannot be made contingent on a worker's realized output. Bonds are allowed to be made contingent only on a worker's decision either to move or to stay at the firm. The more general case, when the bond can depend on y, has proven intractable. Intuition suggests that including this more general case would make it more likely that a separating equilibrium will exist, but if there is enough variability in the job-matching component, θ, then there will be groups of workers in which a pooling equilibrium will still result. The remainder of the paper maintains the assumption that the return on bonds cannot depend on y.

TABLE 3

Mobility and Wages for a Low-Productivity Worker With Bonding

	Period 1	Period 2	Period 3
Output at which a low-productivity worker is indifferent between moving and staying	$y^r(1) = 1.0$	$w_3 - b_2 = 1.0$	—
Fraction of workers who move at end of period	2/3	2/3	—
Wages for a low-productivity worker who never moves	$3/2 - 7/18 = 10/9$	$y = 2$	$y = 2 + 7/9$
Wages for a low-productivity worker who moves only after period one	$3/2 - 7/18 = 10/9$	$w_2 = 56/45$	$y = 2 + 1/5$
Wages for a low-productivity worker who moves after both periods one and two	$3/2 - 7/18 = 10/9$	$w_2 = 56/45$	$w_3 = 6/5$

NOTE: w_2 = the second-period wage for workers who changed jobs after period one; w_3 = the third-period wage for workers who changed jobs after periods one and two.
SOURCE: Author.

move after period one, and two-thirds (one-third) of these workers move again after the second period. Wages for both job-movers and job-stayers increase over the life cycle, although at a slower rate for job-movers.

Notice also that the increase in wages for movers is not monotonic over time: it reaches a maximum in period one and drops off slightly in the last period. Workers who move twice continue to earn more in the last period of their working life than they did in the first period; however, their wages decrease with their last job move. This is consistent with the findings of Bartel and Borjas (1981), who determine that for older men a quit can have either a zero or a negative effect on wage growth.

The example also explains why prior mobility is an indicator of future mobility. The probability that a worker changes jobs in the first period is one-half, while the conditional probability that a worker changes jobs in the second period, given that he changed jobs in the first period, is five-ninths. In contrast, workers who did not move after the initial period will choose never to change jobs. The presence of movers and stayers results because low-productivity workers move more often than high-productivity workers.

The next example illustrates this result by a two-period example. The cost of using a two-period model is that the model can no longer explain why prior mobility is a good indicator of future mobility. The example helps illustrate how these results apply when workers have a continuum of different productivity types.

Example 4: Mobility and Wages in a Two-Period Example With Bonding

The following example is a two-period version of the model presented in example 3. Using the notation defined above, b is the bonus paid in the second period to job-stayers, while w_1 is the first-period wage for all workers and w_2 is the second-period wage for job-changers. In this example, p is allowed to vary continuously with the distribution function, $f(p)$. Thus, each worker has a different productivity level. In addition, we define Λ to be the fraction of workers who change jobs after the first period. Remembering that a worker will change jobs only if $\theta < w_2 - p - b$, Λ is determined as follows:

$$(5) \quad \Lambda = \int G(w_2 - p - b) f(p) dp.$$

The intuition behind this equation is simple. $G(w_2 - p - b)$ is the fraction of the p-productivity workers who change jobs after the first period. This fraction is then multiplied by $f(p)$, the proportion of all workers who have a productivity of p. Summing this product over all productivity types gives the average mobility rate of workers.

The second-period wage for job-movers is determined similarly:

$$(6) \quad w_2 = \int pG(w_2 - p - b) f(p) dp / \Lambda.$$

The intuition behind this equation is similar to that given above. $G(w_2 - p - b) f(p)/\Lambda$ is the fraction of job-movers who have a productivity of p. Multiplying by p and summing over all workers gives the average productivity, or the average output, of a job-changer.

The following example assumes that the matching component, θ, and the individual productivity component, p, are both uniformly distributed: $\theta \sim [-\theta', \theta']$ and $p \sim [p', p'']$. Fol-

TABLE 4

Mobility and Wages for a High-Productivity Worker With Bonding

	Period 1	Period 2	Period 3
Output at which a high-productivity worker is indifferent between moving and staying	$y^r(1) = 1.7$	$w_3 - b_2 = 1.0$	—
Fraction of workers who move at end of period	1/3	1/3	—
Wages for a high-productivity worker who never moves	3/2 - 7/18 = 10/9	$y =$ 2 or 3	$y = 2 + 7/9$ or $3 + 7/9$
Wages for a high-productivity worker who moves only after period one	3/2 - 7/18 = 10/9	$w_2 =$ 56/45	$y = 2 + 1/5$ or $3 + 1/5$
Wages for a high-productivity worker who moves after both periods one and two	3/2 - 7/18 = 10/9	$w_2 =$ 56/45	$w_3 = 6/5$

NOTE: w_2 = the second-period wage for workers who changed jobs after period one; w_3 = the third-period wage for workers who changed jobs after periods one and two.
SOURCE: Author.

lowing example 3, a candidate equilibrium for this example is a pooling equilibrium (where all workers are treated identically ex ante), which maximizes the returns to the highest-productivity worker. Competition for the high-productivity workers, whom firms earn profits by employing, ensures that a pooling equilibrium is obtained by choosing a wage-bonus package (w_1, b) to maximize the expected return of the highest-productivity worker.

$$(7) \quad \max_{w_1, b} \{w_1 + E\max[p'' + \theta + b, w_2]\}$$

such that

$$(1) \quad w_1 + (1-\Lambda)b \leq E(p)$$

$$(2) \quad \Lambda = \int G(w_2 - p - b)f(p)dp$$

$$(3) \quad w_2 = \int pG(w_2 - p - b)f(p)dp/\Lambda.$$

Carlstrom (1989) shows that the problem satisfies

$$b = w_2 - p'$$

If we further assume that p is uniformly distributed between 1 and 2, the corresponding prices and quantities are

$$\Lambda = 1/2 - 1/4\theta',$$
$$b = (3\theta'-2)/(6\theta'-3),$$
$$w_1 = 3/2 - (1-\Lambda)b,$$
$$w_2 = (9\theta'-5)/(6\theta'-3),$$
$$G(w_2 - p - b) = 1/2 - (p-1)/2\theta', \text{ and}$$
$$\Lambda = 1/2 - 1/4\theta'.$$

The above equations indicate that the more disperse θ is (with respect to p), the less important adverse selection is. Increasing θ' raises the wage rate of job-changers and workers' mobility. The reason is straightforward: increasing the variance of θ diminishes the impact of adverse selection, since it increases the incentives for all workers to change jobs. When more workers change jobs, the probability that job-changers are "lemons" is reduced.

Carlstrom also shows that an equilibrium for this example exists if there is enough adverse selection in the labor market, that is, if $\theta' \geq 1$. If we restrict $\theta' = 1$, the corresponding prices and quantities are

$$\Lambda = 1/4,$$
$$b = 1/3,$$
$$w_1 = 5/4,$$
$$w_2 = 4/3, \text{ and}$$
$$G(w_2 - p - b) = 1 - p/2.$$

Notice that the example is consistent with the stylized facts; workers experience a wage increase when they change jobs, yet they earn less over time than job-stayers who earn their output, y, plus their bonus, one-third.

The following section uses this example to discuss questions of optimality.

II. Welfare Implications

Example 4 illustrates another aspect of the model: in equilibrium there is less job mobility than occurs in a world with perfect information. This is not true for all workers, however. High-productivity workers move less often than they would in a world without adverse selection, while low-productivity workers may or may not move less often. There are two reasons for this effect, both of which are due to adverse selection. The

first is identical to that in Akerlof's "lemon" model: adverse selection reduces the future wages for workers when they move and thus reduces the incentive to move. The second effect is due to the posting of bonds in equilibrium, which further reduces the incentives for mobility.

The results of this section are shown with a two-period model, assuming that θ is uniformly distributed. For most of the results, these assumptions can be relaxed. Without bonds, the probability that a worker with a productivity, p, will change jobs is $G(w_2 - p)$; the average probability that a worker changes jobs is $E\{G(w_2 - p)\} = G(w_2 - E(p)) < G(0)$, where $G(0)$ is the probability that a worker would change jobs in a model without adverse selection. The posting of bonds accentuates this effect. In example 4, the unconditional probability that a worker moved was one-fourth, with the lowest-productivity worker moving half of the time, and the highest-productivity worker never moving.

Since mobility is lower in this example than in a model with complete information, it is natural to ask whether a government could increase welfare by subsidizing mobility. An example of such a government subsidy is unemployment insurance. However, since there is no unemployment in the model, unemployment insurance cannot be analyzed. Instead, this paper models unemployment insurance, which decreases the costs of moving, as a subsidy to the wage of job-movers. It therefore asks whether a government can achieve a Pareto improvement by subsidizing the wages of job-movers. Because a government does not have superior information about a worker's productivity, the answer is no.

Subsidizing mobility would not benefit the highest-productivity workers, so taxing them to pay for this subsidy would make them worse off. However, a stronger welfare result can be proven in this model. That is, a government cannot tax first-period wage income to subsidize the wages of job-changers in order to increase aggregate welfare.[5] In fact, it is shown that if a government subsidized the wages of job-movers, there would be no effect on the equilibrium allocations. With a subsidy of s, the equilibrium prices and allocations from the second example are

$$(8) \quad b = w_2 - p',$$

$$(9) \quad \Lambda = \int G(p' - p) f(p) dp,$$

$$(10) \quad w_2 = \int pG(p' - p) f(p) dp / \Lambda \text{, and}$$

$$(11) \quad w_1 = E(p) - (1 - \Lambda) b = E(p) - (1 - \Lambda)(w_2 - p').$$

To verify that subsidizing w_2 by s and taxing first-period income by Λs has no real effect, consider the above equations. Assuming the wage paid to job-movers by firms, w_2, did not change, then from (8) the equilibrium amount of the bonus would increase by s (or bonds would increase by $(1 - \Lambda)s$). In other words, the amount of the bonus paid to the job-stayers would change one-for-one with the subsidy on w_2, leaving mobility the same and thus implying (and verifying the assumption) that the wage paid to job-movers, w_2, remains the same. Therefore, second-period income would increase by s for both movers and stayers, and first-period income would decrease by s. The following are the new equilibrium allocations:

$$(8') \quad b' = w_2 + s - p',$$

$$(9') \quad \Lambda = \int G(w_2 - p - b) f(p) dp = \int G(p' - p) f(p) dp,$$

$$(10') \quad w_2 = \int pG(w_2 - p - b) f(p) dp / \Lambda = \int pG(p' - p) f(p) dp / \Lambda \text{, and}$$

$$(11') \quad w_1 = E(p) - (1 - \Lambda) b' = E(p) - (1 - \Lambda) b - s.$$

Quick inspection of equations (8)–(11) and (8')–(11') shows that subsidizing mobility affects neither mobility, (Λ), nor total wages over time. Mobility stays the same, while wages in the first period for all workers decrease by the subsidy, and net wages in the second period increase by the subsidy. The intuition behind this result is straightforward. Subsidizing mobility benefits the frequent job-movers—the low-productivity workers. In a pooling equilibrium, however, the returns to the highest-productivity workers are maximized. The amount of the bond that would be posted in equilibrium would change one-for-one with the amount of the taxes to eliminate the effects of the government's action.

■ **5** This is in contrast to the welfare implications of Akerlof's model, where a government could subsidize the trading of cars and increase aggregate welfare in the sense that owners of the low-quality cars would gain more than owners of the high-quality cars would lose.

III. Conclusion

Adverse selection is thought to be prevalent in many markets. This paper argues that adverse selection may also be important in the labor market. It can explain why wages tend to increase as workers get older, except for frequent job-movers, whose wages may actually decrease in later years. It also can explain why older workers who move frequently have lower average wages than do infrequent job-changers. Job-movers earn low wages because frequent mobility brands them as low-productivity workers. This effect then decreases the incentives for workers to change jobs.

Thus, adverse selection may seriously impair the ability of workers to change jobs and can interfere with the efficient operation of the labor market. Because of this market failure, it is natural to ask whether a government action to subsidize mobility can reduce the severity of adverse selection and improve the functioning of the labor market. However, it is shown that such a government action will have no real consequences. The reason is that bonds arise in the model in order for firms to compete for the high-productivity workers. Subsidizing mobility hurts the infrequent job-movers (the high-productivity workers), leading firms to increase the amount of bonds required by incoming workers. This increase in bonding offsets the subsidy given to job-movers, leaving the government action ineffective.

The paper also suggests that adverse selection will not be a problem for job-changers if they are paid a piece rate or with a contingent wage contract. Recent actions by firms to pay their workers bonuses and stock options may ease the impact of adverse selection. Future work is needed to address whether these types of contracts are arising as a result of adverse selection and whether these contracts may lead to a more fluid and efficient labor market.

References

Akerlof, George A., "The Market for "Lemons": Qualitative Uncertainty and the Market Mechanism," *Quarterly Journal of Economics,* August 1970, *84,* 488–500.

Bartel, Ann P. and Borjas, George J., "Wage Growth and Job Turnover: An Empirical Analysis," in S. Rosen, ed., *Studies in the Labor Market,* Chicago: University of Chicago Press, 1981.

Becker, Gary S., "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy,* October 1962 (supplement), *5,* 9–49; revised in Gary Becker, *Human Capital,* Second Edition, New York: Columbia University Press, 1975.

Borjas, George J. and Mincer, Jacob, "The Distribution of Earnings Profiles in Longitudinal Data," in Zvi Griliches, ed., *Income Distribution and Economic Inequality,* New York: John Wiley and Sons, 1978.

Carlstrom, Charles, "Three Essays in Labor and Macroeconomics," Unpublished Ph.D. thesis, University of Rochester, 1989.

Feldstein, Martin, "Temporary Layoffs in the Theory of Unemployment," *Journal of Political Economy,* October 1976, *84,* 937–57.

Gottschalk, Peter and Maloney, Tim, "Involuntary Terminations, Unemployment, and Job Matching: A Test of Job Search Theory," *Journal of Labor Economics,* April 1985, *3,* 109–23.

Greenwald, Bruce C., "Adverse Selection in the Labour Market," *Review of Economic Studies,* July 1986, *53,* 325–47.

Harris, Milton and Holmstrom, Bengt, "A Theory of Wage Dynamics," *Review of Economic Studies,* 1982, *49,* 315–33.

Jovanovic, Boyan, (1979a) "Job Matching and the Theory of Turnover," *Journal of Political Economy,* October 1979, *87,* 972–90.

———, (1979b) "Firm-Specific Capital and Turnover," *Journal of Political Economy,* December 1979, *87,* 1246–60.

Leighton, Linda and Mincer, Jacob, "Labor Turnover and Youth Unemployment," in Richard Freeman and David Wise, eds., *The Youth Labor Market Problem: Its Nature, Causes, and Consequences,* Chicago: University of Chicago Press, 1982.

MacDonald, Glenn M., "A Market Equilibrium Theory of Job Assignment and Sequential Accumulation of Information," *American Economic Review,* December 1982, *72,* 1038–55.

Mincer, Jacob, "Movers, Stayers, and Their Wages," Columbia University Discussion Paper No. 67, 1984.

——— **and Jovanovic, Boyan**, "Labor Mobility and Wages," in S. Rosen, ed., *Studies in the Labor Market,* Chicago: University of Chicago Press, 1981.

Oi, Walter I., "Labor as a Quasi-fixed Factor," *Journal of Political Economy,* December 1962, *70,* 538–55.

Riley, John G., "Competitive Signalling," *Journal of Economic Theory,* April 1975, *10,* 174–86.

Rothschild, Michael and Stiglitz, Joseph, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics,* November 1976, *90,* 629–49.

Salop, Joanne and Salop, Steven, "Self-Selection and Turnover in the Labor Market," *Quarterly Journal of Economics,* November 1976, *90,* 619–27.

Spence, Michael, "Job Market Signaling," *Quarterly Journal of Economics,* August 1973, *87,* 355–74.

Topel, Robert and Ward, Michael, "Job Mobility and the Careers of Young Men," manuscript, U.S. Department of Labor, November 1987.

Wilson, Charles, "A Model of Insurance Markets with Incomplete Information," *Journal of Economic Theory,* October 1977, *16,* 167–207.