Though his macro policy views were important to Vickrey, they seem to have been downplayed by the profession when he was awarded the Nobel Prize. His views on macroeconomics were too far from the mainstream to be acceptable. Had he lived, the suppression of his latent ideas would not have succeeded; Vickrey fully intended to use his new found fame as a bully pulpit for his ideas. The thought of this was, we suspect, embarrassing for some in the profession. When, for example, Vickrey gave his presidential address espousing his new ideas to the American Economic Association a few years ago, David Colander remembers seeing a couple of young economists shifting with embarrassment in their seats. One asked the other: “Is this guy for real?” Vickrey was very real and a great economist, who carried out the logic of economic ideas without concern about political sensibility.

Vickrey’s approach to economics was that you must first determine what results are morally justifiable, and then design institutions to achieve them. Currently, most of the profession approaches economics from the opposite direction; they determine what is achievable, and define that as morally justifiable. Of the two approaches, Vickrey’s is, in our minds, the more reasonable. We, and the Eastern Economic Association, will miss him dearly.

REFERENCES


QUIUS CUSTODIET IPSOS CUSTODES?*1

Howard Bodenhorn
Lafayette College

INTRODUCTION

Recent developments in the financial sector have prompted scholars to look to the past to gain a greater understanding of the present. The savings and loan debacle, for instance, prompted a number of studies of early American experiments with deposit insurance, demonstrating that the savings and loan debacle of the 1980s was far from unique; had regulators instead taken a lesson from history, it was avoidable (Calomiris 1989; Wheelock and Kumbhakar 1994; Bodenhorn 1996). Financial innovation, too, has increased scholarly interest in the era of American free banking and other periods characterized by deregulation or rapid technological change (Dwyer 1996; Gorton 1990). A common thread running through these studies is a presumption or an explicit test of the ability of market participants to monitor and delimit the risk-taking activities of financial institutions.

In the wake of events occurring in financial markets in the past two decades, proposals suggesting an increased role for market-based monitoring of financial institutions are being taken seriously. Barth (1991), for example, argues that regulatory oversight alone may not be an effective deterrent to high-risk banking. Market monitoring may be a necessary adjunct as studies by Barth (1991), Kane (1989) and White (1991) suggest that deposit insurance may actually promote inefficiently excessive risk-taking among banks. Deposit insurance is particularly prone to the principal-agent problem because the best interests of the regulatory agency (the agent) and depositors and taxpayers (the principals) often diverge. Providing the appropriate incentives that would entice private market participants to engage in some monitoring activity, according to some of these writers, might mitigate the principal-agent problem by replacing, or at least supplementing, the work of the regulatory agency.

Two decades of research by Joseph Sinkey and his co-authors supports the market monitoring hypothesis while maintaining a role for bureaucratic oversight. Using equity prices, Pettway and Sinkey (1980) demonstrate that equity markets provide regulators with early warning signals of impending failure among large banks. Investor concerns about potential problems at a bank are translated into sharply declining equity prices 33 to 53 weeks prior to a troubled bank’s failure. It is not clear, of course, that an advance warning of nine to twelve months provides sufficient time for regulators to reverse policies or correct problems at troubled billion dollar banks, but it does allow the monetary authority, deposit insurers and others to anticipate the event and begin to plan for impact. Similarly, bank investors reacted rationally and efficiently to the LDC debt crisis of the late 1980s. Banks more heavily invested in LDC debt experienced greater declines in equity prices than banks with lesser exposure (Musumeci and Sinkey 1990a; 1990b). These findings suggest that bank investors, at least, can
and do distinguish between banks based on their relative risk exposures and are not prone to hysteria or panicky divestiture in response to exogenous macroeconomic shocks believed to have deleterious effects on bank portfolios.

Finding that (possibly savvy) investors price shares according to available information on expected profitability or, perhaps, salvage value does not eliminate the possibility that relatively uninformed depositors may initiate a run on a bank. This article explicitly investigates whether agents in early American financial markets effectively monitored financial institutions, punishing those banks that pursued relatively high-risk strategies by more heavily discounting their demandable debt issues without necessarily provoking a run. It employs a methodology similar to that used by Pettway and Sinkey (1988) — the event study — to determine whether market agents were capable of monitoring banks. The results suggest that they were. Widely traded debt (the banknote) issued by failing banks demonstrated a persistent decline in market value commencing about 24 months prior to failure. Early American markets, like modern ones, provided participants with an early warning signal of impending bank failure.

The results presented below complement those found in Gorton [1996]. Employing a sample of banks operating for at least three years prior to failure, I show that market participants could effectively discriminate between solvent and insolvent banks. Gorton, on the other hand, considers a sample of newly opened banks and determines the period necessary for a bank to establish a reputation. Because market agents had little information about the risk preferences of new banks, their debt traded at substantial discounts relative to similar banks that had already established a reputation. Gorton's results suggest that traders quickly differentiated between entrants pursuing low- and high-risk strategies and priced their debt accordingly. Whereas Gorton asked whether young banks whose debt was more highly discounted failed, I ask whether the debt of established banks that ultimately failed was more highly discounted. We are, in a sense, entering the same tunnel from opposite directions.5 Gorton's concern was with reputation formation; mine is with reputation deterioration. In either case, the relevant issue is whether market participants distinguished between low- and high-risk banks and then conveyed their risk assessment of a bank's portfolio through the price at which they stood ready to buy their debt issues.

ANTEBELLUM BANKING SYSTEMS AND BANKNOTE MARKETS

In antebellum America prospective bankers obtained an operating license through either of two procedures. The usual route involved petitioning the state legislature for a corporate charter. This was, understandably, expensive. Obtaining a charter required much legal advice and most petitioners found that the legislative wheels ground slowly without adequate lubrication. Graft and bribery were the rule in many states.6 Nonetheless, it was the petitioners' task to convince the legislature of their integrity and honor (an interesting notion given the widespread use of bribery) and ability to operate a bank and, if successful, the petitioners received a charter granting them the usual corporate privileges as well as the right to issue small denomination promissory notes (banknotes) that circulated as money. In return, the petitioners typically agreed to maintain certain capital and liquidity ratios, to pay dividends to shareholders and taxes to the state, and to submit themselves to the regular scrutiny of various legislative committees or bureaucratic agencies.

With the rise of Jacksonian populism in the 1830s, charter-mongering which soaped of privilege and monopoly fell into disrepute in many states and gave way to so-called free banking.7 Free banking differed from chartered banking in two essential respects. One was that prospective bankers no longer needed a special charter. They had only to satisfy the requirements of a general incorporation law, which typically involved little more than registering with a state regulatory agency a corporate title, an address, and the names of the bank's officers. The other significant departure from chartered banking came in the allocation of the bank's capital. Free banking acts demanded that each bank invest its paid-in capital in state or federal bonds, which were surrendered to the regulators.8 So long as the bank remained both solvent and liquid it received the interest income generated by the bond portfolio. If a bank failed to meet the demands of a single creditor, however, it was considered bankrupt, the bonds were forfeited and sold by regulatory agency to indemnify any unsatisfied creditors once the bank's assets were liquidated. The bonds surrendered to the regulators, then, acted as a form of liability insurance for bank creditors, namely noteholders.

Regardless of whether a bank operated as a chartered or free bank its essential business was the same. Banks took in loans, bills of exchange, commercial paper and other debt instruments, buying them with their own distinct privately-issued money known as banknotes. Privately issued banknotes, like Federal Reserve notes today, were issued in convenient denominations to encourage their use as a medium of exchange. Given the absence of economic alternatives, banks were apparently quite successful in convincing the public to use their notes as currency. The U.S. mint was slow to produce an adequate coinage and neither the federal nor the state governments issued paper money.9 It was bank-issued currency that served as the predominant medium of exchange. Peter Temin [1969, 186-87] estimated that there was about $40 million in specie in the United States in 1820 with more than one-half that amount held in bank vaults, leaving about $20 million in circulation. By comparison, aggregate issues of banknotes, while difficult to estimate for early periods in American history, exceeded $36 million (probably an underestimate) while bank deposits amounted to about $27 million (probably an overestimate) [Pensterman 1965, 60-68]. Clearly, transactions were carried out through the use of bank-supplied currency. Of the $53 million or so in currency chasing goods around the economy, about three-quarters of that currency was bank-supplied and one-half or more was supplied in the form of banknotes.10 Banknotes were the lifeblood of American commerce.

Banknotes circulating within close proximity to the issuing bank exchanged for goods at par because the notes carried an explicit redemption option. That is, any one receiving a banknote received the right to return the note to the issuing bank and demand payment of the face or par value of the note in specie. This redemption option guaranteed that notes circulated at par in the neighborhood of the issuing bank be
cause any departure in the exchange value of the note from par would have yielded arbitrage profits. So long as returning the note to the bank and demanding redemption was essentially costless the par and market value of the note could not long diverge.

As long-distance transactions and traveling carried banknotes away from the issuing bank the exchange value of the note and its par value could diverge, sometimes significantly, depending on the distance or the cost of redeeming the note. That banknote discounts increased with distance from the issuing bank is logical because, as Gorton [1996, 359] points out, pricing a banknote was equivalent to pricing a risky, pure-discount bond whose maturity equaled the time required to return it to its issuer. This is a standard result on the pricing of risky debt: the riskier the bond (banknote), the greater the decline in its value as maturity (distance to issuer) increased [Merton, 1974]. The implication of Merton’s finding, combined with arbitrage possibilities, is twofold: (1) banknotes redeemable at different places will typically trade at different prices at a given location, and (2) all banknotes originating at the same place will trade at the same price at a given location. While there is no direct evidence on the volume of notes moving significant distances away from the issuing bank, there are considerable volumes of indirect and anecdotal evidence suggesting that they did. Fishlow [1964] and North [1974] demonstrated that antebellum America experienced large and growing volumes of interregional trade. Although most interregional trade was financed through bills of exchange, banknotes were almost certainly used to finance some long-distance transactions. One oft-repeated folk tale of American banking is that the use of the term “dixie” in referring to the South can be traced back to New Orleans’ bank issuance of $10 banknotes with the French “dix” for “ten” printed prominently on their face.

Rivermen and traders heading down the Mississippi said they were south-bound in search of “dixies,” which they then carried back and spent in Kentucky, Ohio and western Pennsylvania. Less fanciful, perhaps, but more direct evidence of the wandering nature of banknotes is that country banks often maintained redemption accounts with city banks. As previously noted, as banknotes moved away from the issuing bank they depreciated in exchange value. Depreciated banknotes were less acceptable in exchange because exercising the redemption option was more costly. To mitigate the depreciation and ensure a wide circulation of their notes many country banks developed ties with city correspondents. City correspondents agreed to redeem the country bank’s notes at par so long as the country bank maintained an adequate balance. In New England a single bank — the renowned Suffolk Bank — established a consolidated, regionwide redemption system [Whitney, 1878]. In New York, the legislature mandated that country banks establish note redemption agreements in New York City or Albany. The Pennsylvania experience, however, demonstrates that redemption systems required neither an endogenously generated hierarchical system (the Suffolk system) nor legislative fiat. Pennsylvania’s country banks voluntarily sought out city correspondents and negotiated mutually beneficial redemption arrangements [Bodenborn, 1990]. The very existence of these redemption systems suggests that notes often trav-
Market-based monitoring presumes the existence of an effective monitoring and disciplining device. Calomiris and Kahn [1991] argue that demandable debt (deposits or banknotes) provides just such a device. As Fama [1980, S84] noted, the priority and maturity structure of a firm's debt-equity ratio may be adopted to provide the appropriate incentives to the monitors. In modern corporations, for instance, uncollateralized loans are usually lower priority than the fixed-payoff promises offered suppliers of raw materials and labor. Because the priority structure places uncollateralized creditors at greatest risk, they have the greatest incentive to monitor. Higher priority debt holders and residual claimants, like stockholders, have little incentive to monitor and effectively delegate the task of liquidating an insolvent firm to low priority debt holders.

As Calomiris and Kahn [1991, 500] point out, an analogous priority structure is captured by banks with demandable-debt contracts coupled with a first-come, first-served constraint. Monitors exclusively bear the costs of screening, but are compensated for their services by moving to the head of the queue should a bank's liquidation be favored. By moving to the rear of the queue, a bank substitutes full compensation upon closing out, but they accept that possibility because the monitors bear the costs and guarantee non-monitors a positive payout. Monitors will call for liquidation before a bank's net worth is fully depleted, thereby guaranteeing themselves full compensation and all others at least partial remuneration. Because the monitors have placed a fraction of their wealth at risk, a bank's continued existence provides a credible signal about the monitors' assessment to uninformed debt holders.

But Ramakrishnan and Thaker [1984] and Fama [1990] suggest that monitors need not necessarily place themselves directly in harm's way to be credible. Pure information producers may arise who disseminate valuable information about a debt whose ultimate value cannot be known with certainty ex ante and can only be estimated with error. Although such rating agencies do not place their own wealth at direct risk in that they hold debt instruments, they establish themselves as credible monitors if incorrect current ratings decrease the market value of future ratings. And this is more likely if doubts arise about the motives of the rating agency or the cause of the erroneous rating. So long as ratings are viewed as reasonably accurate and uninformed principals believe the market will punish consistently inaccurate ratings, there is little incentive to duplicate their monitoring activities (Fama, 1980, S87). Ramakrishnan and Thaker [1984, 416], however, believe that competition among monitors is beneficial. Competition may mitigate moral hazard problems because multiple monitors police one another as well as the principals. History appears to support Ramakrishnan and Thaker's contention because multiple agents have frequently appeared to monitor the activities of firms with fiduciary responsibilities. In antebellum America, for instance, two distinct types of agents appeared who monitored and delimited the banks' activities.

In antebellum America specialized rating systems and agents arose to monitor banks. Banknote brokers were independent businessmen who collected and processed information on banks and stood ready to buy the notes of most banks. These brokers made a market in banknotes and profited from arbitrage possibilities much like modern foreign exchange dealers. Broker purchased notes at a discount from face value and returned them to the issuing bank for redemption into specie. As Bedenham [1988] and Gorton [1986] have shown, the prices established in these informal secondary markets reflected the cost of carrying the notes back to the issuing bank and a risk premium reflecting the subjective probability that a bank might fail between the time of purchase and its expected redemption date.

As Ramakrishnan and Thaker [1984] predicted, most cities had several brokers buying banknotes in these informal secondary markets. The traditional interpretation was that, with a small initial capital requirement, the industry was competitive (or contestable). It is obvious, however, that given the positive marginal costs of gathering additional information, direct social costs would have been minimized had a single broker operated in each city. But the potential for moral hazard—a propensity for a single information producer to provide unreliable information once a credible reputation was created—generated the incentives for several independent brokers to operate simultaneously and monitor one another as well as the banks. As the number of brokers increased, expected moral hazard or misinformed costs decreased and, in many cities, the number of independent brokers grew rather large. A New York City directory for 1842, for instance, listed 51 exchange offices or money collectors whose principal business was buying and selling banknotes [Longworth, 1842]. A New Orleans city directory for 1843 listed no fewer than 30 banknote brokers [New Orleans Annual, 1843]. Because of contemporary trade flows, these cities may have had more brokers than most, but most substantial cities had two or more banknote brokers.

Note brokers can be seen as a special type of Diamond's [1984] delegated monitor who operated under Calomiris and Kahn's [1991] first-come, first-served constraint. Efficient pricing of a bank's note required information gathering and processing. Note brokers, then, were in a favorable position to recognize the early warning signs of imminent default and exercise their cashing out option before the general public. It was in this regard that the value of competition came to the fore. With many monitors, a single broker receiving an erroneous or randomly noisy signal was not in a position to inefficiently liquidate a solvent bank. Liquidation of any particular bank required a critical mass of brokers interpreting the available information similarly. A single, nervous broker simply would not be in a position to force a costly and inefficient liquidation based on an erroneous signal.

Though note brokers monitored, their information remained private, reducing the information content of their actions. Lacking the widespread dissemination of prices, the monitoring problem was, once again, pushed back one step. Instead of monitoring the banks directly, note holders and depositors would have had to moni-
For the brokers by observing prices paid for various notes. But given the number of note-issuing banks and the number of note-buying brokers in many cities, monitoring the brokers would have been nearly as costly as monitoring the banks themselves. This cost was reduced by the emergence of the second set of important players in the banknote market, namely, the banknote reporters.

Updated and published weekly, banknote reporters were tabloid-style newspapers that provided a whole range of information useful to merchants and manufacturers. These reporters usually offered brief articles on market trends, reports from correspondents in distant markets, and several columns of wholesale prices. The real purpose of the banknote reporter, however, was to provide businessmen and the public with information about banknotes. Reporters, of course, specialized in providing current prices as determined in secondary markets as well as descriptions of counterfeiters, usually identifying them by bank and denomination. A businessman offered an unfamiliar note could turn to a recent reporter and determine the likelihood of the note's authenticity and its current price. Like note brokerage, banknote reporting was a competitive business [Billstein, 1949]. A prominent commercial city might have two or more specialized reporters. General audience newspapers, too, often printed banknote prices along with other business news in a "Market News" or similar section, not unlike the abbreviated exchange rate tables found in business sections of most modern daily newspapers.

Antebellum financial markets therefore provide the possibility of testing the view that markets endogenously generate effective monitors and information producers in a sector typically thought to be subject to sharp information asymmetries. If these markets operated efficiently, market prices of banknotes should have reflected the perceived probability of a bank's default by a number of independent monitors. To test this hypothesis, publicly reported banknote discounts for a sample of banks failing between January 1838 and December 1843 were collected from Bichnell's Counterfeit Detector, Bank Note Reporter and General Price Current published in Philadelphia. The test interval (1838-1843) was chosen for a number of reasons. First, it was a relatively compact period in which a number of banks throughout the country failed, yielding sufficient observations to draw some generalizations. Second, it provides a test of the monitoring hypothesis when monitoring would have been most difficult and risky — the financial panic of the late 1830s and the depression of the early 1840s. Under such macroeconomic conditions, market-based monitoring was certainly put to the severest of tests.

Banknote reporters published extensive lists of all ongoing and failed banks and listed a discount from face value at which brokers currently purchased each bank's notes. A typical entry would take the form: Bank of Gotham ... Gotham City ... 1d, meaning that notes of the Bank of Gotham were currently purchased by brokers at a 1 percent discount. Because these discounts reflected both redemption (or transportation) costs and risk discounts, the value of each bank's notes was standardized, eliminating redemption costs and leaving only a pure risk discount. This risk discount was captured by measuring the excess discount for a failing bank's note compared to the modal discount for all banks located in the same state as the failing bank. This excess discount was calculated as:

\[
\text{Excess Discount} = \frac{\text{failing bank discount} - \text{modal discount}}{100 - \text{modal discount}}.
\]

The advantage of this measure, as Gorton [1996, 367] points out, is that it captures the discount on failing banks normalized for the modal discount. And, as previously noted, the modal discount is the relevant basis of comparison because, in the long run, the discount on the notes of all banks at a particular location will tend toward the mode. Changes in the modal discount, as well, will reflect changes in such wide ranging phenomena as, say, state-level bank regulation or exogenous macroeconomic shocks. Gorton [1996, 367] argues that the modal discount accurately reflects the influence of a variety of risk factors on banknote prices and is, therefore, a robust benchmark.

Because banks failed throughout the period 1838 to 1843, a common dating mechanism as well as a standardizing mechanism was needed. The month of failure for each failing bank was assigned time \( t = 0 \) and the calculation of the excess discount was made for each of 26 failed banks for 24 months prior to and 12 months subsequent to the failure announcement. If monitors produced relatively accurate information and reacted to it efficiently, the excess discount on a failing bank's notes should have begun to decline several months prior to failure. Though it should not necessarily be used as a basis of comparison, recall that Pettaway and Sinkey [1980] found that equity prices of troubled banks began to decline nine to twelve months prior to failure.
Figure 2 presents the excess discount for the sample of failing banks. Fully 24 months prior to failure, the average failing bank's notes traded at a discount about 20 percent greater than the modal discount. Twelve months prior to failure, the excess discount had increased to about 40 percent; and by four months prior to failure, the excess discount had increased to about 60 percent of the mode. The path followed by the average failing bank's excess discount using a quadratic regression specification is presented in Figure 1 as the "Estimated" curve. A plausible interpretation is that troubled banks were identified by note brokers and other market participants at least 24 months prior to failure. In response to the deterioration in the bank's reputation, its management embarked upon some policies to reestablish itself as a legitimate going concern, witnessed by the modest decline in the excess discount on the bank's notes in the 18 to 24 months prior to failure. Despite attempts to correct problems, banknote market participants became increasingly skeptical of the bank's righting itself as the excess discount increased at an increasing rate until the bank finally succumbed.

This result is not unexpected given the results reported in Pettaway and Sinkey [1985] who found that, among a sample of banks failing in the 1970s, bank equity prices began to decline nine to twelve month prior to failure. A century earlier bank debt prices began to decline about 24 months prior to failure. While the two worlds were radically different and any number of confounding influences were likely to have affected banking and financial markets, it is interesting to note that differences in the lead time of recognizing problems at troubled banks are broadly consistent with our expectations. Fama [1990] suggested that monitoring activity would be influenced by one's standing in the priority structure and that debt holders would engage in more monitoring than shareholders. While the results should not be taken as more than suggestive, they are in general agreement with Fama's contention.

The question remains whether or to what extent brokers and reporters may have represented part of the problem in that their mistakes may have fored otherwise solvent banks into bankruptcy. It was suggested earlier that competitive brokerage and reporting markets mitigated the possibility that erroneous (or erroneously interpreted) signals could lead to the inefficient liquidation of an otherwise sound institution. But banknotes, after all, were fundamentally different from stocks or bonds. A sharp decline in the price of a bank's stock may provide a great deal of relevant information to investors about the goings-on within the firm that allow for reorganization without necessarily bringing about the firm's liquidation. A sharp decline in the market price of a bank's notes, on the other hand, may induce most note holders to redeem the notes and, given a fractional reserve banking system, necessarily force its liquidation. The interpretive problem is that without a great deal of complementary information concerning the portfolio of the failed banks it is impossible to distinguish between those banks which got what they had coming and those forced out of business because banknote brokers and reporters were mistaken.

Using banknote prices it is possible to construct a test of the hypothesis that participants in banknote markets could monitor accurately, punishing excessively risky banks and rewarding others. If banks responded to downgraded ratings by altering their portfolios to allay the concerns of the raters, it should be possible to uncover cases where a bank's notes went to a high excess discount and later returned to the modal discount. Figure 2 presents the average excess discount of 30 banks for which brokers downgraded their ratings and then survived for at least four years after the initial downgrade. It is readily seen that the excess discount on these banks' notes rose sharply at the time of the initial rating downgrade, rising to nearly 30 percent. But it also sharply declined, falling to an average excess discount only one-half the initial excess discount in the second month after the downgrade. By seven months after the initial downgrade the average excess discount had fallen to just 5 percent. Of the 30 non-failing banks whose notes sold at substantial excess discount, 25 saw their notes return to the modal value in 15 months or less. The other five remained a point of concern for note brokers and witnessed a sharp secondary decline nine months after the initial downgrade. Of these five, one's notes returned to the modal discount only after 32 months, another at 30 months, with the three remaining banks' notes still trading at modest excess discounts 48 months after the initial rating downgrade.

Of course, there are two interpretations of these cases. One would be that brokers correctly identified risky banks and, through the discipline of prices, forced them to shape up. The other would be that brokers made mistakes for which otherwise well-run banks paid the price. In order to allay customer concerns they may have had to hold less-than-optimal portfolios, accept below-market interest rates on loans, pay above-market rates on deposits, hold unprofitably large reserves or some combination thereof. While the latter interpretation cannot be unequivocally ruled out, the
findings presented earlier as well as the complementary findings in Bodenhorn [1993] and Gorton [1996] suggest that these antebellum financial markets priced bank debt efficiently.

TESTING THE EFFICIENCY HYPOTHESIS

If the market for banknotes was, in fact, reasonably efficient, more sophisticated tests of the monitoring hypothesis are possible. Because banknotes were widely held debt instruments traded and priced in centralized (if informal) markets, we can apply the common event study methodology to determine if markets reacted quickly and efficiently to new information about a bank’s long-run viability.

In general, the most important influences on a particular bank’s note price would be general macroeconomic movements that would affect the expected profitability of the financial sector generally. But because the object of interest is the change in the value of a bank’s notes due to changes in firm-specific information, it is necessary to purge movements in a bank’s note price from movements resulting from economy-wide shocks. The resulting residual term provides a measure (albeit imperfect, depending on the efficiency of the market and the appropriateness of the model) of the perceptions of informed traders about the viability of a bank. Deriving the residual return requires purging from the gross return movements in the market as a whole.

Assuming that the return on a given bank’s notes \( R_{t, j} = (P_{t, j} - P_{t-1, j}) / P_{t-1, j} \) was principally a function of the returns to banknotes generally, we can estimate a model of the form specified in Equation (2), where \( R_{t, j} \) represents the returns to a broad index of banknotes.\(^{25}\) For those familiar with the finance literature they will quickly recognize this as a variant of the so-called market model, which assumes that the return to a particular firm’s equity is related to movements in the equity market generally.

\[
R_{t, j} = \alpha_{j} + \beta_{j} R_{m, t} + \epsilon_{t, j} \]

\[
\epsilon_{t, j} \sim N(0, \sigma_{\epsilon}^{2}) \quad \text{and} \quad \text{cov}(\epsilon_{t, j}, \epsilon_{t, s}) = 0 \quad t \neq s
\]  

Equation (2) is an appropriate specification if the process generating the expected returns follows a stationary, multivariate normal distribution (Eckbo, 1963, 301). As with standard equity pricing models, \( \beta_{j} \) captures systematic co-movements of the return on bank’s notes with the return on a market portfolio of banknotes. The serially uncorrelated, mean zero error term captures the impact of firm-specific or region-specific information events and random price changes. The coefficients \( \alpha_{j} \) and \( \beta_{j} \) were estimated using OLS regressions utilizing returns from month \( t = 42 \) through \( t = 25 \) for each of 18 banks for which returns were available at least 42 months prior to failure \( (t = 0) \).\(^{26}\) Within the period \( 24 < t < 12 \), abnormal residual returns were calculated as \( \epsilon_{t, j} = R_{t, j} - \hat{a}_{j} - \hat{\beta}_{j} R_{m, t} \), where \( \hat{a}_{j} \) and \( \hat{\beta}_{j} \) are the OLS estimates of \( a_{j} \) and \( \beta_{j} \).

The relevant statistics are denoted \( CAR_{t} \), which is the monthly abnormal residual returns cumulated over a period of interest for bank \( i \), and \( CAMR \), which is the cross-sectional average of mean daily returns cumulated over a relevant interval. Figure 3 presents the cumulative cross-sectional mean daily residual return \( (CAR_{t}) \) for the sample of failing banks. The results reinforce the conclusions drawn from the simpler test above, except that the returns on notes of soon-to-fail banks do not show signs of continued significant decline until six months prior to failure. Standard tests of significance fail to show a statistically significant decline in cumulative cross-sectional note returns over the longer periods \( 24 < t < 0 \) (\( t = -0.55 \)) or \( -12 < t < 0 \) (\( t = -1.22 \)).\(^{25}\) In the last six months before failure \( (6 < t < 1) \), however, the \( CAMR \) demonstrates a statistically significant decline \( (t = -1.87) \) at the five-percent level against a null hypothesis that the expected return is less than zero. That the cumulative abnormal mean returns fail to demonstrate a significant decline more than six months before failure is not particularly troubling because Macklinlay [1997, 28-34] demonstrates that the power of the test declines sharply as the period of cumulation increases, as the time separating observations increases, or as the number of securities decreases. A combination of a relatively long period of interest (24 months), monthly observations, and a relatively small sample all conspire against a strongly statistically significant estimate. Given that the likelihood of finding a statistically significant result was low, finding a significant decline in the six months preceding failure should only increase our confidence in the result.

Considering the cumulative individual returns \( (CAR_{i}) \) for individual banks confirms the belief that the returns to holding a soon-to-fail bank’s notes were negative. Over the interval between 24 months prior to failure \( (24 < t < 1) \), 13 of 18 banks experienced negative residual returns of which 10 were significantly negative at the five-percent level. In the twelve months prior to failure, the returns on holding notes
of 14 banks declined, of which 12 were significant. Similarly, in the six months prior to failure (-6 ≤ i < -1), 13 banks' notes experienced negative cumulative returns of which 9 were significantly negative.

Evidence from antebellum banknote markets, then, suggests that market participants were capable of monitoring financial institutions holding highly idiosyncratic portfolios. This confounds, but does not contradict, a popular view about the special problems surrounding bank regulation and oversight. Because banks and other financial intermediaries arise to exploit their ability to intermediate between imperfectly informed lenders and investors prone to moral hazard (over-reporting the expected value of their investment projects) in a world of asymmetric information, the market finds it difficult to value these idiosyncratic portfolios. Evidence from the antebellum era suggests that this view may underestimate the capabilities of market-based monitoring. It appears that specialized antebellum banknote market participants monitored banks and did it rather well. Of the 26 bank failures considered here, only one surprised note brokers. Notes of the Planters & Merchants Bank of Mobile traded at the modal Alabama note price until the news of the bank's failure leaked to the public. Insiders escaped from this kind of completely unanticipated event that justifies non-market regulatory oversight and insurance. Unanticipated shocks, like the failure of the Planters & Merchants, may be just the sort of event that creates an environment of distrust, makes debt holders jittery and provokes costly and uncalled-for bank runs.

CONCLUDING REMARKS

Previous studies of antebellum financial markets have focused almost exclusively on the costs of multiple note issues. Cagan (1965, 20), for example, argued that:

... the overwhelming variety of circulating notes confused the public and abetted fraud. ... The direct cost to society of transacting business with this confusing mass of paper, "bunk, gruezy, issued by nobody knows whom," can be measured by the expense of supplying and using banknote detectors. ... and no gains offset these costs, which therefore burdened the economy as a whole.

Even those with some sympathy for the operation of free markets in the supply of money and credit have tended to focus on the costs of multiple note issues [Friedman and Schwartz, 1963, 22-23; Rockoff, 1975; Rolnick and Webor, 1984; Dryer, 1996]. While most have shown the costs to be small, none has explicitly addressed the possibility that society received some offsetting benefits from this confusing mass of paper. When equity markets were thin and the stock of many banks closely held by compact groups of merchants or by single families, equity markets were unlikely to provide adequate monitoring of a bank's risk-taking. Where equity markets could not, banknote markets provided an essential monitoring service. Market specialists arose who priced bank debt in informal secondary markets and their activities provided valuable information to otherwise uninformed traders. Banknote reporters, operating in a competitive environment, provided timely information about the brokers' activities, which allowed the public to make informed decisions about the quality of the currency offered in exchange. The very presence of these markets and monitors may also have mitigated the ever-present possibility of a bank run.

Cagan's observation that multiple note issue involved significant social costs is undoubtedly correct (relative, at least, to the modern standard of a uniform national currency), it is not clear that the use of the next-best antebellum alternative — specie — would have been welfare enhancing. Most of the specie available for use in transactions in early America had been minted in foreign countries and was, therefore, not denominated in dollars. Additionally, much of this foreign specie was old, worn, abraded and clipped. A combination of a non-dollar denominated and an excessive weight of foreign currency would have resulted in weighing almost by every coin and haggling over every exchange. This alternative was not costless and early Americans may have been quite rational in their preference for a not-so-confusing mass of paper money.

While lessons for modern life can sometimes be drawn from historical experiences, we must avoid drawing overly broad and inappropriate generalizations. It is unlikely, of course, that lessons learned from the history of multiple note issues can have much direct relevance to current bank policy. It is notable, however, that several studies of the savings and loan crisis suggested that market-based reform is called for. Most of these calls to reintroduce market-based monitoring do not propose to replace the current system of regulation, oversight and deposit insurance with a pure laissez-faire banking system. Instead, the proposals suggest market-based reforms designed to supplement traditional regulation. Suggestions are varied, but have included such things as market value accounting, private reinsurers, true co-insurance, risk-adjusted capital requirements, and the use of subordinated debt [Petajeway and Simkey, 1989; Barth, 1991; Dowd, 1992]. The presumption underlying these suggestions is that markets provide effective monitoring above and beyond what regulators can provide. If the antebellum experience can serve as a guide, these presumptions may have some foundation.

NOTES

I would like to thank Richard Grossman, Michael Huxport, Eugene White, three anonymous referees and the editor of this Journal for many constructive comments that substantially improved the quality of this effort. They should not be implicated for any remaining errors of fact or interpretation.

1. Who should watch the watchmen?
2. In a recent essay on American free banking Dryer (1996, 4) identifies a short piece discussing the possibility that electronic money may become popular and remov[e] the nineteenth century banknote. This may be seen as a very real possibility, especially with the growth in offshore banking and electronic transactions.
3. I thank an anonymous referee for pointing out this metaphor. Our findings are complementary for a number of other reasons as well. Most important, our data come from different sources published in the same city. That our findings are generally consistent suggests that the market was an efficient

QUIS CUSTODES IPSOS CUSTODES?
generator of information, which participants reported reasonably accurately. There has long been a concern that tests of efficiency in banknote markets were tests of the accuracy of a single source. This study, along with Garrett's, suggests that participants were reporting information developed in a competitive marketplace.

4. The literature on early American banking is replete with anecdotes of legislative gridlock and widespread bribery. Hummer [1987] reported that the usually covert negotiations became so openly egregious in 1813 that New York's governor felt compelled to precipitate the assembly in order to clear the draft.

5. Rockoff [1979] remains one of the best descriptions of the rise of free banking.

6. To ensure that the banks were deposited prior to a bank's commencing operations, the oversight agency was responsible for printing the bank's notes and retained control over the engraved plates.

7. There was, in fact, a constitutional ban on state issues of paper money, but various states at various times found avenues around the ban.

8. Many of the deposits held in banks in the early nineteenth century were not demand or checkable deposits. This would increase the banknote-currency ratio.

9. Garrett [1980] notes that redemption costs and distance to the issuing bank was positively corre- lated throughout the antebellum era.

10. Speeds of these propositions are given in Garrett [1980, 387-93].

11. Bills of exchange were debt instruments issued principally by wholesale merchants in long-distance trade. A New York cotton factor, for example, might buy cotton in Louisiana, ultimately destined for a Massachusetts textile mill. Instead of taking the cotton for cash in New Orleans, the factor transferred a bill payable in 90 to 90 days in New York. In the interim, the cotton was shipped, sold to the textile mill with the funds necessary to retire the bill on deposit in a New York City bank. See Bodenhamer [1989] for a more complete description of the breadth and depth of exchange markets in antebellum America.

12. Modern examples of these kinds of monitors include such organizations as Standard & Poor's or Moody's bond rating agencies, TEW consumer credit ratings, as well as a number of others. 

13. Examples of antebellum rating agencies will be discussed in the next section.

14. To put the proposition more formally, suppose that the price of a debt instrument can be modelled as 

\[ P = f(A, x) \]

where \( P \) is observable by expending an amount \( C \) and \( x \) is a white noise, random variable. Suppose one monitors believe that erroneous ratings result from movements in expected prices, which will remain credible. If, on the other hand, a belief deviates that erroneous ratings result from failures to expend \( C \) and observes \( P \), the value of the monitor's ratings will decrease and their profits will fall.

15. Technical efficiency, of course, requires that the sum of direct information gathering and indirect (expected moral hazard) costs be minimized. Estimating whether that point was reached is difficult, but a well-known result is that firms in competitive markets operate in a technically efficient manner.

16. Philadelphia, for example, had access to extensive lists of banknote prices in two specialized reports — Rockoff's Counterfeit Detector and Bank Note Reporter and Van Court's Counterfeit Detector and Bank Note List — as well as the general market news paper — Philadelphia Free Press — and less extensive summaries in general readership newspapers.

17. It is important to define a "failing" bank. The sample included in the study of failed banks only if it closed in a relatively short and imposed losses on noteholders or shareholders. A bank that was voluntarily liquidated and paid its creditors in full was not considered a failure; these banks were considered "closed." Rockoff's apparently drew a similar distinction. Some non-failing banks were listed as "closed," some were identified as "failed" or "broken." Only those banks listed as "failed" or "broken" were included in the sample.

18. Certainly, more than 20 banks failed between 1833 and 1843, but only those banks for which the month of failure could be determined with some confidence were included. For most banks, these dates were determined as the first month in which a bank was listed as "failed" in Rockoff's. Failure dates for New York free banks were taken from Bullock and Weber [1984, 284-85].

19. The estimating equation took the form:

\[ P = \beta_0 + \beta_1 P_{-1} + \beta_2 C + \epsilon \]

where \( P_{-1} \) and \( C \) represent period of commission, \( C \) is the number of days of commission, and \( \epsilon \) is the number of days of commission in the underlying regression equation. See Gardee, Silber and White [1982] or MacKinley [1987] for a more thorough derivation.

REFERENCES


Biddulph's Counterfeit Detector, Bank Note List and General Price Current. Various issues, 1825-1845.


Cagan, P. The First Fifty Years of the National Banking System — An Historical Appraisal, in Banking and Monetary Studies, edited by Darre Coven, Homewood, Ill.: Irwin, 1969, 15-42.


AN ANALYSIS OF ADAM SMITH'S THEORY OF CHARITY AND THE PROBLEMS OF THE POOR

Thomas D. Birch
University of New Hampshire at Manchester

INTRODUCTION

A well-established literature recognizes Adam Smith's concern with issues of distributive economic justice (Winch, 1978, 91-9; Heilbroner, 1982; Young and Gordon, 1996), particularly the plight of the poor (Himmelweit, 1985, 46-62), and some related research into how Smith's thought might justify a supportive role for the state in providing assistance to the poor (Baum, 1992). However, the existing literature has failed to analyze the potential role of private charity in addressing the problem of poverty using Smith's writings as a foundation. As a result, the case for government intervention on behalf of the poor within the context of Smith's thought has not been fully explored. This paper addresses this shortcoming by examining Smith's theory of charity as it relates to the poor. Taken as a whole, Smith's writings reveal a complex theory of charity which allows for both self-regarding and altruistic motives of donors, assigns a prominent role to recipient behavior as a determinant of charitable giving, and recognizes psychological, ethical, historical, social, as well as economic factors influencing the nature and level of charitable activity.

In The Theory of Moral Sentiments, (hereafter TMS) Smith often speaks of the virtue of beneficence (which includes acts of charity, kindness, love, friendship and the like), but this virtue has received comparatively little attention among economists, evidently due to two reasons. First, the core theoretical argument in The Wealth of Nations (hereafter WN) would appear to be that economic altruism is largely unnecessary in a competitive market system inhabited by just and prudent individuals. Second, beneficence is described by Smith in TMS as "the ornament which embellishes" society rather than being a virtue essential to social order or economic growth (1776, 78). The comparative neglect of beneficent acts such as charitable giving is not entirely justified, however, even within the context of Smith's own thought. As argued elsewhere (Baum, 1992), Smith was concerned about imbalances of economic power prejudicial to the poor created by the mercantilist system which he, argued, largely benefited the capitalist class of merchants and manufacturers. Reinforcing this perspective is the fact that Smith emphatically declared that the "obvious and simple system of natural liberty" which he advocated in WN was a utopian vision which "not only the prejudices of the publick, but what is more uncontroulable, the private interests of many individuals, irresistibly oppose" (WN, 687, 471). The issue of charitable giving to benefit the poor is therefore highly relevant in a "second best" world created by rent-seeking capitalists or other market failures.