International Journal of Forecasting 28 (2012) 145–160

Contents lists available at SciVerse ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting and explaining aggregate consumer credit delinquency behaviour

Jonathan Crook *, John Banasik

Credit Research Centre, The University of Edinburgh Business School, 29 Buccleuch Place, Edinburgh EH8 9JS, United Kingdom

ARTICLE INFO

Keywords:
Finance
Co-integration
ARIMA models
Error correction models
Time series
Unit roots

ABSTRACT

We model aggregate delinquency behaviour for consumer credit (including credit card loans and other consumer loans) and for residential real estate loans using data up until 2008. We test for cointegrating relationships and then estimate short run error correction models. We find evidence to support the portfolio explanations of declines in credit quality for consumer and for real estate loans, but support for the reduced stigma explanation was restricted to real estate loans. Evidence supportive of household-level explanations of irrational borrowing and unexpected net income shocks was found for consumer and real estate loans, but evidence of strategic default was restricted to the volume of consumer loans and real estate loans, and not for credit cards. We also found that the error correction model gave forecasts of the volume of delinquent consumer debt which were of an accuracy comparable to that of an ARIMA model.

1. Introduction

The recent rise in consumer loan defaults and mortgage defaults in the US and Europe has emphasised the significance of accurate credit risk modelling and the interdependence between the banking sector and the real economy. The recent crisis had many causes, but one of them was the increase in default rates of sub-prime mortgage loans in the US. A contributing factor of this was the rapid extension of loans to high risk borrowers whose ability to repay was highly dependent on the state of the macroeconomy. When house price inflation began to fall and interest rates, fuel prices and eventually unemployment increased, many of these borrowers defaulted (Arner, 2009; Crouhy, Jarrow, & Turnbull, 2008). Although this considerable increase in default rates has occurred only since 2006, there are good reasons for believing that the state of the macroeconomy has more long run effects on the proportion of borrowers that default in any one year.

It is important for lenders to be able to explain and predict the aggregate consumer delinquency over time. An increase in the total consumer delinquency, *ceteris paribus*, may increase the need to raise interest rate margins to compensate for the increased risk, and also in order to retain sufficient liquidity. A significant increase in delinquencies may cause lenders with low capital adequacy ratios to become insolvent, causing widespread failures by contagion. The Basel II Accord (BIS, 2006) allows banks to determine their own capital requirements by using their own models to forecast future probabilities of default. These probabilities must be 'through the cycle' (probabilities that do not vary with the business cycle), and a common method of obtaining these is to use a technique that involves modelling default rates in terms of macroeconomic variables (see Heitfield, 2005).

In this paper, we model aggregate consumer default rates in the US over a twenty year period, including the period of their recent escalation. We do this for consumer loans and mortgages separately. We use a cointegration technique to explain the long run relationships between

* Corresponding author. Tel.: +44 131 650 3802; fax: +44 131 650 8337.
E-mail address: j.crook@ed.ac.uk (J. Crook).

0169-2070/$ – see front matter © 2011 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.ijforecast.2010.12.002

default rates and the macroeconomy, and model changes in the default rates in terms of deviations from the long run relationship and short-run changes in the macroeconomy. We also compare the predictive performances of these models with the performances of ARIMA models, in order to see which methodology gives more accurate forecasts using the especially challenging task of forecasting recent events. We find evidence to support the existence of long run 'equilibrium' relationships between the level of interest rates and the level of debt outstanding on the one hand and aggregate default rates on the other, but, surprisingly, not between the level of house prices and the level of default rates. However, we *do* find that *changes* in house prices have a significant effect on *changes* in default rates, as do changes in disposable income, unemployment, consumer confidence, and interest rates. We also find that the two forecasting methodologies both give highly accurate forecasts of default rates, and are about equally accurate. Had these models been available in mid-2008, the subsequent increases in default rates could have been forecast accurately. We make two contributions. First, we offer a model of aggregate consumer default rates for the US using cointegration techniques, and thus separate long run 'equilibrium' relationships from short run dynamic relationships. Second, we compare the forecasting performances of this econometric technique and ARIMA. To the best of our knowledge, this has not been done before.

The next section of the paper reviews the related literature, and we subsequently explain our model. In view of the especially high interest in mortgage defaults, we present our results for these sectors in separate sections. The final section concludes.

2. Related studies

The literature suggests that there are essentially three reasons why a borrower defaults on a loan. First, a borrower may manage his/her finances poorly, due to hyperbolic discounting, leading to a preference for 'irrational' immediate expenditure (Laibson, Repetto, & Tobacman, 2003). Second, there is an 'ability to pay' hypothesis, namely that a borrower will fail to pay on time when an income or expenditure shock occurs that was not expected at the time when the loan was taken out. The causes of such shocks include unexpected job loss, marital breakdown, family bereavement, health problems, increases in interest payments on loans, and so on. Third, there is the 'strategic default hypothesis', where, when a loan is used to buy a real asset (such as a house), and if the capital market is perfect, with no transaction costs or reputation effects, a borrower would increase his wealth if he defaulted on a loan when the value of it was greater than the value of the asset (Kau, Keenan, & Kim, 1993, 1994).[1] When considering aggregate default rates over time in the United States specifically, several explanations have been advanced. Observing the increase in credit card delinquency rates between 1994 and 1997, Gross and Souleles (2002) propose two explanations. First, that the proportion of borrowers that were high risk has increased, and these are the borrowers who have defaulted. Second, that borrowers 'have become more willing to default', given their risk characteristics, because the social stigma of default and the associated loss of a future credit supply have declined.

Previous empirical studies that have related the delinquency of credit card debt and mortgages to macroeconomic variables have used either duration models or time series models. We start with duration models for credit cards. All three such studies of which we are aware have estimated *account* level duration models using macroeconomic variables as time varying covariates. Gross and Souleles (2002) used a panel of over 200,000 credit card borrowers. Surprisingly, they found that none of the unemployment rate in the county of residence, the per capita income or house prices in the region was significantly related to delinquency, and, together with measures of borrower risk, they could only explain a small proportion of the changes in delinquency rates over time. The residual was tentatively ascribed to the trend of reduced stigma. However, FCIC data suggests that if the period under consideration is extended to between 1992 and 2006, the delinquency rate on credit card debt was trended downwards, if anything, and the same was true of the total consumer debt. Agarwal and Liu (2003) also used panel data for credit card holders for 1995–2001, and found that the probability of a credit card holder missing three consecutive payments in a particular period, given the card holder's predicted level of risk, was increased if the lagged unemployment rate in the county or state of residence was higher, but that the actual change in the unemployment rate had no effect. The account balance three months earlier also had a positive effect on the hazard rate. Bellotti and Crook (2009) estimated a proportional hazards model for a sample of credit cards issued by a UK bank between 1997 and 2001, and found that the base interest rate, real earnings, production and house prices all had a significant effect on the hazard rate.

Turning to account level panel models of duration for mortgage debt delinquency, Lambrecht et al. (1997) used a survival model applied to 5272 borrowers in the UK and found evidence which was more in favour of the ability to pay argument than of the strategic default hypothesis. However, none of the variables which they included varied over time. Deng, Quigley, and Van Order (2000) estimated a competing risks model of prepayment and default, for mortgages granted between 1976 and 1983, to investigate

[1] More realistically, if transaction costs do exist and default does reduce the chance of a borrower gaining future loans, the option to default will not be exercised until the debt is somewhat greater than the asset value, since defaulting removes the option of either defaulting or repaying in the future (Kau et al., 1994). Lambrecht, Perraudin, and Satchell (1997) point out that the costs of default are higher for some than for others. For example, those for whom access to debt is particularly important will experience a higher cost if defaulting reduces the chance of borrowing in the future. According to the permanent income hypothesis, these are individuals who expect their income to rise in the future (Deaton, 1992). Note also that, unlike a Chapter 7 bankruptcy declaration in the United States, a default in some countries, such as the UK, does not prevent creditors from pursuing the debtor for repayment. In such countries, this latter point removes the reason for strategic default.

the extent to which the hazard rate can be explained by the strategic default hypothesis. For default to be optimal in the presence of transaction costs, the put option must be in the money, and trigger events such as divorce or unexpected unemployment must occur. The time varying annual divorce rate and quarterly unemployment rate in the state of residence were both found to have a significant effect on the probability of default, as was the probability that the put option was in the money. Teo (2004) used a sample of mortgage loans in Singapore to test an eclectic range of hypothesised determinants of the hazard rate. He found that whilst neither the characteristics of the property bought nor those of the borrower could explain the rate, those of the mortgage and of the macroeconomy could. Teo's evidence may be interpreted as supporting both the ability to pay and strategic default hypotheses. However, Teo's study is limited by its small sample size (657 cases), and by collinearity.

Whilst account level duration models account for borrower specific characteristics, these studies model delinquency over relatively short periods of time and do not cover an entire business cycle, with the possible exception of the work of Bellotti and Crook (2009). It is therefore questionable whether the variation in the macroeconomic variables over time is sufficient to estimate their effect accurately. Furthermore, these studies have not considered the autocorrelation properties of their model residuals.

In contrast, a small number of time series studies have considered data on *aggregate* default rates, which, while they do omit borrower specific characteristics, cover much longer periods. One of the earliest of these studies was performed by Sullivan (1987). Using data from 1975 to 1986, she found evidence supporting the ability to pay hypothesis, and showed that the willingness of banks to lend affected default rates. However, Sullivan did not consider some important factors which one would expect to be important in the explanation of delinquency rates, such as the level of interest rates, and there is evidence that her empirical model may have been misspecified. Grieb, Hegji, and Jones (2001) empirically modelled bank card delinquency rates over the period 1981–1999. They found that they were explained by debt to income ratios, which were taken to represent the capacity to pay, with no evidence that delinquency was due to job market conditions, high interest rates or high credit supply. They do, however, find evidence that borrowers defaulted on credit card debt before other types of consumer debt. However, the empirical model in their study has a low explanatory power, and omits the possibility that an error correction mechanism may be estimated and may be more informative than the model chosen.

Two papers have used error correction models to model mortgage delinquency, using data for the UK and England and Wales respectively. Whitley, Windram, and Cox (2004) found that the proportion of mortgage loans which are at least six months in arrears is related to the mortgage income gearing, unemployment, and the loan to value ratio for first time buyers. However, the lack of regression diagnostics, the imputation of quarterly data from semi-annual data, and the lack of explanation of the structure of

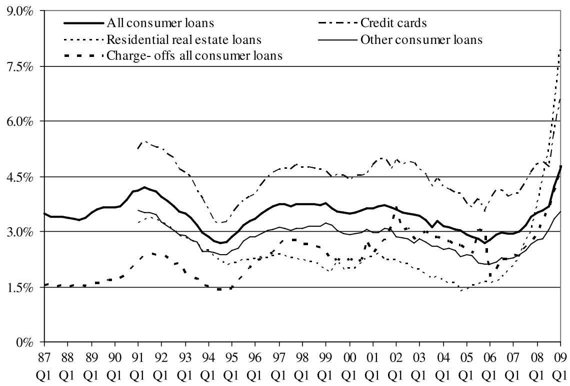

Fig. 1. US seasonally adjusted delinquency rates (30+ days overdue) and charge-off rates for different loan types.

their model limits the usefulness of these results. Figuera, Glen, and Nellis (2005) use quarterly data for 1993–2001 and find that the proportion of loans that are three months overdue is related to the unemployment rate, the loan to income ratio for first time buyers, the unwithdrawn equity, and the debt service ratio.

Clearly, explanations of why some individuals default and others do not may be able to explain aggregate default rates. For example, if there is an increase in the incidence of catastrophic net income shocks, irrational borrowing or negative equity, then one would expect an increase in the aggregate default rates.

Overall, the literature suggests that any variations over time in the aggregate delinquency rates for *unsecured* credit are due to variations in both the ability of the average borrower to make repayments and the risk distribution of borrowers due to bank lending policies. For *secured* lending, one can also add variations in the values of real assets relative to the debt outstanding on them. However, apart from Bellotti and Crook (2009), none of these papers test the forecasting ability of their models, and none of the studies of US default rates provide a thorough treatment of the time series properties of their data. We now turn to the observed patterns in US household delinquency.

3. Patterns in delinquency and charge offs

Fig. 1 plots the delinquency and charge off rates as a percentage of the outstanding debt[2] for all consumer loans and mortgages extended by all US commercial banks from 1987 until 2009. During this period, the trend in charge off rates is distinctly upwards, whereas that of 30+ days delinquency is slightly downwards until 2006. This suggests that, until 2006, the average period of time which was taken before a delinquent loan was charged off was gradually shortened, especially over the period 1997–2002. The values in 2002 Q1, where the charge off

2 The delinquency rate is the value of loans 30+ days overdue as a percentage of the debt outstanding at the end of the quarter; the charge off rate is "the value of loans removed from the books and charged against loss reserves …measured net of recoveries as a percentage of average loans and annualized" (FRB).

rate slightly exceeds the delinquency rate, are probably due to a slightly different method of calculating the two rates, and possibly different methods of applying seasonal adjustment.[3] The sudden rise from 2006 is readily apparent. We construct a model to explain these patterns of aggregate delinquency in Section 4.

The delinquency rates for all consumer loans in Fig. 1 mask different patterns in the rates for different types of loans. Note that consumer loans consist of credit card loans plus other consumer loans, with residential real estate loans being separate. The trend for all three types of loans was downward from 1992 to 2006 (2005 for real estate loans) and rose rapidly after that. However, the delinquency rate for real estate loans appears to have been affected very little by the business cycle trough in late 1994, whilst the rate for consumer loans was substantially affected, and credit card loans especially so. Perhaps surprisingly, the consumer loans seem to be positively correlated in the mid 1990s. That is, as the real disposable income declined until 1995 Q4 and rose thereafter, the default rates on consumer loans declined as well, though they stopped mirroring income after about mid-1997. One possible explanation for this is that as the level of income falls so does the demand for debt, and thus the less creditworthy find that repayments relative to income decline and they are less likely to miss a payment or to stay overdue. If there is a critical level of debt outstanding above which there is a disproportionate number of defaulters, then when income declines, the overdue debt will decline faster than the debt outstanding. One would expect this to apply especially to short term debt (consumer debt, and especially credit card debt) rather than to debt where the borrower expects to repay over many years (residential debt). Of course, a detailed examination of these possible explanations requires us to examine the time series properties of the series, and construct a multivariate model, which we do in the next section.

4. The model

We can think of the movement of the aggregate volume of debt between different states over time. We could represent this movement in a conventional transition matrix as shown in Table 1, where the states are: 1 = No credit, 2 = Up to date, 3 = 30+ days overdue and 4 = Charged off, and $v_{i,j}$ is the volume of credit which moves from state i in period t to state j in period $t+1$. We are not assuming that $v_{j,j}$ remains constant over time. Let the period of time denoted by t be one quarter. Certain values of $v_{i,j}$ must necessarily take on the value of zero. These are v_{12}, v_{13}, v_{14}, v_{41}, v_{42}, v_{43} and v_{44}.

Table 1
Repayments transition matrix.

	1	2	3	4
1	v_{11}	v_{12}	v_{13}	v_{14}
2	v_{21}	v_{22}	v_{23}	v_{24}
3	v_{31}	v_{32}	v_{33}	v_{34}
4	v_{41}	v_{42}	v_{43}	v_{44}

The change in the stock of overdue debt consists of v_{23}, which is the volume which moves from being up to date to being 30+ overdue; v_{31} and v_{32}, which are the volumes which move from 30+ overdue to no credit or up to date, respectively; and v_{34}, which represents the volume which moves from 30+ to being charged off. Letting $d_t = v_{23}$, $p_t = (v_{31} + v_{32})$ and $c_t = v_{34}$, we can write:

$$s_t - s_{t-1} = (d_t - p_t) - c_t, \tag{1}$$

where s_t = real volume of consumer debt which is 30+ days overdue in quarter t.

We model delinquency rates in terms of the ability to pay hypothesis, and, for loans for residential real estate, we also include a variable to represent the strategic default hypothesis. Thus, we assume that the volume of debt which is 30+ days overdue at the end of a quarter is correlated with the level of nominal interest rates (*ri*), the volume of debt outstanding (*ccout*), personal disposable income (*pdi*), and expectations about future income over the period. The interest rate and the level of disposable income affect the ability of a borrower to repay, and thus also the aggregate number of borrowers who default. Expectations of a higher future income may lead a borrower to wish to borrow more both now and in the future, and thus he will not wish to risk his ability to do this by missing payments now. For real estate loans we also include the level of real house prices (*rhp*), the argument being that the lower house prices are, controlling for the level of debt outstanding, the greater the proportion of borrowers for whom the value of the debt will exceed the value of the property plus transaction costs, and the greater the advantage of default, assuming that the lender does not continue to pursue the debtor.

These arguments imply that the change in the stock of overdue debt, the level of $(d_t - p_t) - c_t$, is correlated with changes in these explanatory variables. An increase in interest rates or a reduction in disposable income, which, at the level of a borrower, could be the result of a catastrophe such as job loss or marital break-up, when aggregated across borrowers, would be expected to result in an increase in the aggregate volume of overdue debt.

We assume that the long-run relationship between the stock of overdue debt and its determinants is linear, and therefore write

$$s_t = \delta + \boldsymbol{\delta}^T \mathbf{x}_t + \varepsilon_t, \tag{2}$$

where $\mathbf{x}$ is a vector of covariates and δ and $\boldsymbol{\delta}$ are a scalar and a matrix of parameters to be estimated, respectively. ε_t represents an error term.

4.1. Estimation

The vector error correction representation of Eq. (2) is

$$\Delta s_t = \boldsymbol{\beta}^T \Delta \mathbf{x}_{t-1} + \theta(s_{t-1} - \delta_1 - \boldsymbol{\delta}_2^T \mathbf{x}_{t-1}) + \varepsilon_t, \tag{3}$$

where $\boldsymbol{\beta}$ and θ are a matrix and a scalar respectively, and are to be estimated. Engle and Granger (1987)

[3] The delinquency rates were seasonally adjusted by the authors using X12. The charge off figures were adjusted by the FRB.

Table 2
Phillips-Perron unit root tests.

	Levels (with trend)	Adjusted t-statistic	Differences (without trend)	Adjusted t-statistic
Consumer delinquency types				
Real bank consumer credit total	Lnrdelsa	−1.067	dlnrdelsa	−5.569**
Bank credit card	Lnccsa	−1.143	dlnccsa	−5.763**
Other bank consumer credit	Lnosa	−1.198	dlnosa	−5.154**
Mortgage loan	Lnrnsa	3.462	dlnrnsa	−3.914**
Explanatory variables				
Real real estate credit outstanding	Lnrnoutsa	−0.971		
Real consumer credit outstanding	Lnrccoutsa	−1.979	dlnrccoutsa	−6.579**
Personal loan interest rate	Lninsa	−2.782	dlninsa	−8.607**
Consumer sentiment index	Lnsent	−1.958	dlnsent	−12.984**
Real personal disposable income	Lnrpdisa	−2.125	dlnpdisa	−13.134**
Real house price index	Lnrhpisa	1.666	dlnrhpisa	−4.250**
Real real estate credit outstanding	Lnrnoutsa	−0.971	dlnrnoutsa	−7.213**
Mortgage interest rate	Lnmisa	−3.578*	dlnmisa	−9.848**
Real total household debt	Lnrtdoutsa	−1.466	dlnrtdoutsa	−3.494*
Credit card interest rate	Lnccintsa	−2.494	dlnccintsa	−6.964**

Test period: 1987 Q1–2009 Q1 for all variables except for the consumer sentiment index, for which it is 1987 Q1–2008 Q4, because 2009 Q1 data was not available. In all cases, the bandwidth is 4 (Newey-West using the Bartlett kernel).
* Significant at the 5% level, one-sided test (MacKinnon).
** Significant at the 1% level, one-sided test (MacKinnon).

showed that if variables in the $\mathbf{x}_t$ vector and s_t are integrated of order 1, and if a cointegrating vector exists, then there is a vector error correction representation of the model, of which Eq. (3) is an example, where ε_t is white noise. The expression in brackets in Eq. (3), the error correction mechanism, represents the deviation of s_t from its long-run value of $\delta_1 + \boldsymbol{\delta}_2^T\mathbf{x}_{t-1}$. Eq. (3) could either be rewritten and estimated as an autoregressive distributed lag model or estimated as a vector error correction model (VEC), and in principle the two sets of estimated structural parameters should be the same (Patterson, 2000). Because it revealed more information overtly, we chose to estimate the VEC form. We therefore tested all of the variables for the order of integration, and finding them to be $I(1)$, except for the mortgage interest rate, we proceeded to estimate the long-run relationship using the Johansen cointegrating ML procedure and then to estimate the ECM representation (Johansen, 1988).

In general, having estimated the cointegrating relationship using Eq. (3) with several lags, we then estimated the short-run dynamic model:

$$\begin{aligned}\Delta s_t &= \alpha + \sum_{l=1}^{4}\beta_{0l}\Delta s_{t-l} \\ &+ \sum_{l=0}^{4}\beta_{1l}\Delta ri_{t-l} + \sum_{l=0}^{4}\beta_{2l}\Delta pdi_{t-l} \\ &+ \sum_{l=0}^{4}\beta_{3l}\Delta ccout_{t-l} + \sum_{l=0}^{4}\beta_{4l}\Delta\,(\text{optimism})_{t-l} \\ &+ \sum_{l=0}^{4}\beta_{5l}\Delta(uet)_{t-l} + \sum_{l=0}^{4}\beta_{6l}\Delta(hp)_{t-l} \\ &+ \theta_1\left[s_{t-1} - \delta_1 - \delta_2 ri_{t-1} - \delta_4 ccout_{t-1}\right. \\ &\left. - \delta_5\,(\text{optimism})_{t-1} - \delta_6(rhp)_{t-l}\right]. \end{aligned} \tag{4}$$

Here we assumed that the variables in the $\mathbf{x}_t$ vector were weakly exogenous and thus included as Δx_t terms. To allow more distant changes to affect the short-run dynamics of the model, we included the first differences of the ri, pdi, $ccout$, nominal house prices (hp), the unemployment rate (uet) and optimism variables, lagged up to four quarters and then tested down to a parsimonious form. The variables in the cointegrating vector were selected to accord with reasonable *a priori* predictions. These were that (a) it would seem implausible that delinquency would be lower at higher levels of interest rates; and (b) higher levels of consumer debt outstanding would result in higher delinquency. Disposable income is omitted from Eq. (4) for reasons which we give in the next section.

5. Results

The data for the *volume* of overdue debt on consumer loans to commercial banks was estimated from the delinquency rates published on the FRB website. For total consumer loans, the delinquency rate was multiplied by the volume of consumer loans, both seasonally unadjusted, and then was seasonally adjusted using the Stats Canada X12 routine. All of the variables were seasonally adjusted using X12 unless only seasonally adjusted values were available. The natural logs of all variables were then used. Unfortunately, due to a lack of data on the corresponding *amounts* of debt outstanding on credit cards, residential mortgages or other consumer loans, our dependent variable for these types of loans is the delinquency *rate*. We could not find interest rates for each separate type of loan for the entire period of our data. However, we were able to find data for credit card interest rates, mortgage interest rates and the mean rate for 24 month personal loans.

We first checked whether the variables were stationary using a Phillips-Perron test. We assumed a time trend for levels but not for first differences. The results are shown in Table 2. From this it can be seen that all of the variables were integrated of order 1, except for the mortgage rate, and thus their first differences were stationary. The results

Table 3
Johansen cointegration tests.

H_0:		Trace statistic	5% cv	Max-eigenvalue statistic	5% cv
	Consumer credit Total real default volume (Lnrdelsa)				
$r = 0$		75.65	63.88	36.10	32.12
$r \leq 1$		39.56	42.92	20.39	25.82
$r \leq 2$		19.17	25.87	16.69	19.39
$r \leq 3$		2.48	12.52	2.48	12.52
	Lags in ECM = 4				
	Credit card default rate (Lnccsa)				
$r = 0$		66.28	63.88	37.02	32.12
$r \leq 1$		29.27	42.92	14.43	25.82
$r \leq 2$		14.83	25.87	10.24	19.39
$r \leq 3$		4.59	12.52	4.59	12.52
	Lags in ECM = 4				
	Other loans default rate (Lnosa)				
$r = 0$		80.52	63.88	44.21	32.12
$r \leq 1$		36.31	42.92	16.02	25.82
$r \leq 2$		20.29	25.87	13.83	19.39
$r \leq 3$		6.46	12.52	6.46	12.52
	Lags in ECM = 4				
	Residential real estate loans default rate (Lnrnsa)				
$r = 0$		119.78	88.80	41.01	38.33
$r \leq 1$		78.77	63.88	35.45	32.11
$r \leq 2$		43.32	42.92	22.71	25.82
$r \leq 3$		20.61	25.87	10.91	19.39
$r \leq 4$		9.71	12.52	9.71	12.52
	Lags in ECM = 4				

The estimation samples for the models were: volume of consumer credit: 1988 Q2–2008 Q1; for delinquency rates for credit cards, other consumer loans and residential real estate loans: 1992 Q2–2009 Q1.

of the Phillips-Perron test for the mortgage interest rate varied according to the time period of data that was used. For example, it suggested that the mortgage rate was $I(0)$ if one used the entire data series available to us (up to 2009 Q1). However, if we used just up until 2008 Q1, the test suggested that the mortgage interest rate was $I(1)$. We chose to include the mortgage interest rate in the ECM, since the tests which we use for a cointegrating vector would indicate no vector if the mortgage rate was not integrated of order 1. The cointegrating vector and the short run dynamic models were estimated using the following data periods: volume of consumer credit: 1988 Q2–2008 Q1; default rates for credit cards, other consumer loans and loans on real estate: 1992 Q2–2008 Q1. The difference in the beginning date is due to data availability. We omitted 2008 Q2–2009 Q1 from the estimation sample so that we could use it to assess the forecasting accuracy of the model.

Because different hypotheses may explain delinquency for different types of loans (credit card loans, other loans and loans on real estate), we considered the delinquency behaviour for each type of loan separately.

5.1. Volume of consumer credit

Table 3 shows the results of the Johansen cointegration tests. For the volume of delinquent consumer debt, we excluded the disposable income from the long-run model, because when it was included, either the trace and eigenvalue statistics rejected the null hypothesis that there exists at least one cointegrating vector, or the elasticities on the income or other variables were implausible. The top panel shows both the trace and maximum eigenvalue statistics for the included variables, and they reject the hypothesis of at most zero cointegrating relationships but not the hypothesis that there is at most one. We therefore conclude that there is only one cointegrating relationship. Column 2 of Table 4 shows this relationship normalised on the *volume* of delinquent debt. Since the values are all in logs (except for the trend), the coefficients can be interpreted as elasticities. The relationship shows that in the long-run, *ceteris paribus*, the higher the nominal personal loan interest rate and/or the volume of consumer debt, the greater the volume of consumer debt that is 30+ days overdue. The asymptotic t-statistics suggest that both are statistically significant. The delinquency elasticity of the volume of debt in equilibrium (at 2.48) is somewhat lower than for nominal interest rates (at 3.42).

Columns 2 and 3 of Table 5 show the short-run dynamic equation after variables have been removed on the basis of t-statistics and *a priori* expectations, and assuming that all of the independent variables are weakly exogenous in explaining the volume of delinquent debt. The error correction term is highly significant and negative, meaning that the greater the amount by which the volume of debt in default exceeds its long-run value in one quarter, the larger the decrease in delinquent debt in the next quarter, which is consistent with our expectations. The value implies that

Table 4
Cointegrating vectors (normalized).

Dependent variable (delinquency rates or volume)		Consumer credit			Residential real estate rate (lnrnsa)	Residential real estate debt (lnrnoutsa)
		Total real volume (lnrdelsa)	Credit cards rate (lnccsa)	Other rate (lnosa)		
Estimation period:		1988(2)–2008(1)	1992(2)–2008(1)	1992(2)– 2008(1)	1992(2)–2008(1)	1992(2)– 2008(1)
Independent variables						
Personal loan interest rate	lninsa	3.421179 (5.138)**		0.224736 (0.895)		
Credit card interest rate	lnccintsa		1.052104 (2.440)**			
Mortgage interest rate	lnminsa				4.440855 (4.814)**	−0.058910 (−0.379)
Real consumer credit outstanding	lnrccoutsa	2.476919 (6.560)**				
Real total household debt outstanding	Lnrtdoutsa		3.293867 (3.703)**	0.326231 (0.850)		
Real house price index	lnrhpisa				−1.287682 (−0.828)	0.819381 (3.128)**
Consumer sentiment index	lnsent	0.250140 (0.772)	4.149131 (5.550)**	2.637235 (7.538)**	−1.217450 (−0.579)	0.005718 (−0.016)
	Trend	0.000646 (0.353)	−0.042407 (−3.046)**	−0.005393 (−0.860)	0.057898 (3.135)**	0.012822 (4.123)**
	Constant	−32.96678	−71.12901	−17.22614	2.02118	2.298972

Asymptotic t-statistics are in parentheses.
* Significant at the 5% level.
** Significant at the 1% level.

Table 5
Short run dynamic equations.

Dependent variable:	dlnrdelsa		dlnccsa		dlnosa	
Estimation period:	1988(2)–2008(1)		1992(2)–2008(1)		1992(2)–2008(1)	
	Coefficient	*t*-stat	Coefficient	*t*-stat	Coefficient	*t*-stat
Independent variable						
Constant	0.006963	1.904	0.063755	4.034**	0.036956	5.693**
Δdependent variable						
ddepvar(−1)	0.294305	3.269**	0.277482	2.544*	0.292584	3.222**
ddepvar(−2)			0.409872	4.321**		
ddepvar(−3)	−0.178716	−2.039*			0.272562	2.868**
ddepvar(−4)	0.191735	2.850**				
Δpersonal loan interest rate						
dlninsa	0.247647	2.061*				
dlninsa(−1)	−0.318031	−2.230*				
dlninsa(−2)					0.398035	2.371*
dlninsa(−4)	−0.283033	−1.928				
Δcredit card interest rate						
dlnccisa(−1)			0.995690	4.382**		
dlnccisa(−2)			−0.459833	−1.847		
Δreal cons. credit outstanding						
dlnrccoutsa	1.769506	11.424**				
dlnrccoutsa(−1)	−0.863338	−4.234**				
dlnrccoutsa(−3)	0.332669	1.731				
Δreal total debt outstanding						
dlnrtdoutsa(−4)			−2.711368	−3.059**		
Δreal personal disp. inc.						
dlnrpdisa					−0.960631	−2.799**
dlnrpdisa(−3)			−1.178360	−2.282*		
dlnrpdisa(−4)			−1.348418	−2.646*		
Δ(log)optimism						
dlnsent	−0.167160	−2.520*				
dlnsent(−1)			−0.580658	−3.570**		
dlnsent(−2)	−0.122058	−1.753			−0.195921	−1.702
dlnsent(−3)	−0.152598	−2.070*			−0.271455	−2.294*
dlnsent(−4)	0.141314	1.981	0.288269	1.979		
Δunemployment rate						
dlnuets(−2)			0.310217	2.519*		
dlnuets(−4)	0.124044	1.981				
Δnominal house prices						
dlnhpisa					−2.281425	−5.463**
dlnhpisa(−1)	−0.743795	−2.887**				
Error correction						
ecmlnrdelsa1(−1)	−0.136846	−6.206**				
ecmlnccsa24(−1)			−0.137630	−4.031**		
ecmlnosa5(−1)					−0.155482	−4.988**
Adjusted R^2	0.798193		0.524616		0.540635	
DW	2.295758		2.111985		2.094834	
Durbin's *h* alt.	−0.104661		−0.635365		−0.772350	
Jarque-Bera $\chi^2(2)$	6.731451*		0.112595		6.360840*	
RESET2 $\chi^2(1)$	0.404732		1.117542		0.004008	
LM het. test $\chi^2(1)$	0.053380		0.614157		0.191852	
F-statistic	20.528973**		7.320408**		10.268245**	

All variable changes are in logs.
The total volume of delinquent consumer debt (lnrdelsa), consumer credit outstanding (lnrccoutsa), personal disposable income (lnrpdisa), and total household debt outstanding (lnrtdoutsa) are all in real terms. The house price index (lnhpisa) is in nominal values.
* Significant at the 5% level.
** Significant at the 1% level.

only 13.7% of the deviation of delinquent debt from its equilibrium value is removed in the next period.

Considering aggregate explanations, these results are consistent with the credit quality explanation (increases in credit volume tend to be gained by accepting higher risk borrowers), but lend little support to the stigma hypothesis. Considering the long run relationship, one would expect a positive marginal effect of the volume of consumer debt outstanding (conditional on the personal loan interest rate) if the default rate was constant, but the elasticity of 2.48 indicates that the volume of delinquent debt would increase at a faster rate than the volume of consumer debt, implying an increase in the default rate with an increase in consumer debt. The credit quality argument is also sup-

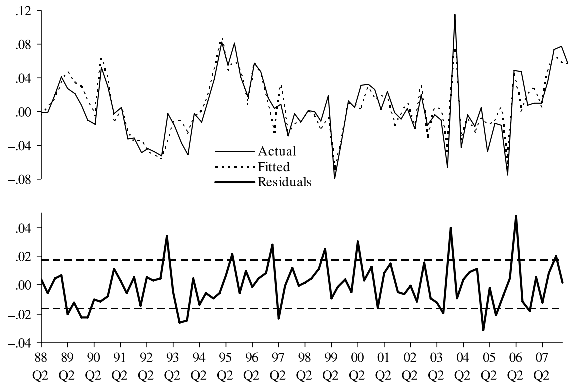

Fig. 2. Changes in the log volume of delinquent credit.

ported by the positive effect of the personal loan rate, because the proportion of applicants that are high risk is expected to be higher at higher interest rates, due to adverse selection (Stiglitz & Weiss, 1981). The insignificance of the trend variable is not consistent with the stigma hypothesis, and the short run results further support this interpretation. The greater the increase in consumer debt, the greater the increase in delinquent volume.

Turning to explanations at the household level, the results support all three hypotheses. The hypothesis of irrationality (irrationally borrowing debt that cannot be repaid) is supported by the magnitude of the elasticity on the level of debt outstanding and the positive sign on the change in debt in the short run equation. It is also supported by the negative sign on the lagged interest rate term. This suggests that the greater the decrease in the rate, the greater the volume of delinquent debt one quarter later, which is consistent with households reacting to the decline in the rate irrationally, and taking on more debt than they can repay, possibly because of hyperbolic discounting. Support for the adverse shock hypothesis is given by (1) the positive sign on the personal loan interest rate: at higher levels of this rate, the volume of delinquent debt is higher; and (2) the elasticity on the volume of debt. Further support is given by the short run dynamic models, where we found that a greater increase in the loan interest rate results in a greater increase in delinquency volume in the same quarter, and also that the larger the decrease in households' expected financial situations relative to their current situations, the greater the increase in delinquency volume. The strategic hypothesis is also consistent with the sign on the short run equation of the change in nominal house prices: the greater the fall in house prices, the greater the increase in delinquent debt.

Fig. 2 shows the observed and predicted changes in volumes of overdue consumer debt. In-sample, the model predicts relatively poorly in the first quarter of 1993, the fourth quarter of 2003 and the second quarter of 2006, when the predicted values were much smaller than those observed in all three cases; and also in quarters 3 and 4 of 1993, when the predicted values were much larger than those observed. We now turn to the delinquency rates.

5.2. *Delinquency rates*

We modelled the delinquency rates for two types of consumer loans separately: credit card loans and other consumer loans. These together make up the total consumer loans — the variable corresponding to the volume of delinquent consumer debt in the previous section. Due to data restrictions, we were unable to model the volume of delinquent debt in each category. Instead, the dependent variables were the volume of debt 30+ days overdue as a percentage of the end-of-quarter debt outstanding. The model used the assumptions above.

5.2.1. *Credit card delinquency rates*

Panel 2 of Table 3 shows the results of the Johansen cointegration tests for the credit card delinquency rate. We excluded the real disposable income, because when we experimented with it we either found no cointegrating relationship or found that the implied elasticities on income or other variables were implausibly high or of an implausible sign. Both the trace statistic and the max eigenvalue test suggest that we can reject the null of no cointegrating vectors, but not the null that at most one vector exists. We conclude that there is one vector, and the parameters of the vector, normalised by the delinquency rate, are shown in column 3 of Table 4.

Considering aggregate explanations first, the long run relationships again support the credit quality explanation. The mean credit card interest rate and the volume of the total household debt outstanding are both positively related to the default rate. There is no support for the stigma hypothesis, with the effect of the trend (conditional on the interest rate and debt outstanding) being negative.

Looking at household level explanations, the results support two of the hypotheses, but not the strategic default hypothesis. The irrational behaviour hypothesis gains support from the large and positive elasticity on the household debt outstanding in the long run equation and the negative sign on the lagged interest rate in the short run equation, with the latter indicating that the greater the reduction in the interest rate, the greater the increase in the delinquency rate two quarters later. The adverse shock hypothesis is supported by the positive effect of interest rates in the long run and the one-period lagged positive effect of a change in credit card interest rates on the increase in the default rate in the following quarter. The lagged effect of increased unemployment also is consistent with the adverse shock explanation: the greater the increase in the unemployment rate in one quarter, the greater the increase in delinquency two quarters later. Similarly, the lagged effect of an increase in income resulting in a decrease in delinquency rate nine months later is also consistent, although the effect takes rather a long time. The argument that households miss a payment on their credit card because they have a negative net equity in their house receives no support, since the effect of house prices was insignificant in both the long run and short run models. This is entirely plausible, since one would expect this hypothesis to apply only to secured lending.

The size of the adjustment coefficient on the cointegrating vector, −0.138, is similar to that for the volume of the consumer debt equation.

5.2.2. Other consumer debt delinquency rates

For other consumer loans, the interest rate used was the 24-month personal loan interest rate. The cointegration tests are shown in panel 3 of Table 3, and agree that there is one cointegrating relationship. Following normalisation on the delinquency rate, the estimated cointegrating vector for delinquency rates on other consumer loans is given in column 4 of Table 4.

Neither of the aggregate explanations gains support from the long run equations. Of the household level explanations, support for the adverse shock hypothesis is provided by the lagged positive sign on the personal loan rate and the negative sign on the real disposable income. The negative sign on real house prices is consistent with the strategic default hypothesis, and makes sense if asset prices move consistently, so that house prices reflect the value of assets bought with these loans.

6. Residential real estate loans

When estimating the cointegrating vector for residential loans, we included a fixed rate mortgage interest rate, because the vast majority of first lien primary mortgages are fixed rate. For example, Buck, Kinneckel, and Moore (2006), using the Survey of Consumer Finance, found that only 15% of those with a first lien primary mortgage had one with an adjustable interest rate in 2004 and only 11% in 2001. We experimented with the inclusion of real personal disposable income, but when included, it yielded implausible signs or elasticities on income or other variables, or few variables that were significant. We subsequently obtained two cointegrating relationships, as shown in panel 4 of Table 3. We normalise the first on the delinquency rate and the second on the residential real estate debt, and obtain the results shown in column 5 of Table 4.

These results support the stigma explanation of Gross and Souleles (2002) for delinquency: conditional on the mortgage interest rate, real house prices and sentiment, the trend in delinquency was upwards over the 1990s and 2000s. Evidence in favour of the credit quality argument is provided by the strong positive effect of the level of the mortgage rate on the level of delinquency in the long run equation and the positive effect of the increase in the mortgage rate on the increase in the delinquency rate in the short run equation. Since we use the *fixed* interest rate, changes in this are unlikely to affect current borrowers, but they would affect new borrowers, who, if they are offered and accept relatively high rates, may subsequently find that they are less able to repay than were borrowers who accepted lower rates. In short, poorer quality applicants have been accepted, with banks charging higher margins to cover increased risk. The positive conditional trend effect is also consistent with a credit quality effect. Notice that our results relate to a long period of time, not merely to the period of the recent crisis.

All three household level hypotheses are supported. The adverse shock hypothesis is supported by the effect of changes in disposable income on the changes in delinquency from the short run equation (Table 6). The irrationality hypothesis is supported by the positive sign on the level of the mortgage rate and on the increase in the mortgage rate in the short run equation. Because we are using the fixed mortgage interest rate, it is likely that a change in this rate would not affect a significant proportion of current borrowers, but it would mean that new borrowers were accepting higher rates than previous borrowers and then missing payment(s). This is consistent with the irrationality hypothesis. The strategic default hypothesis is supported by the short run negative effect of lagged changes in house prices.

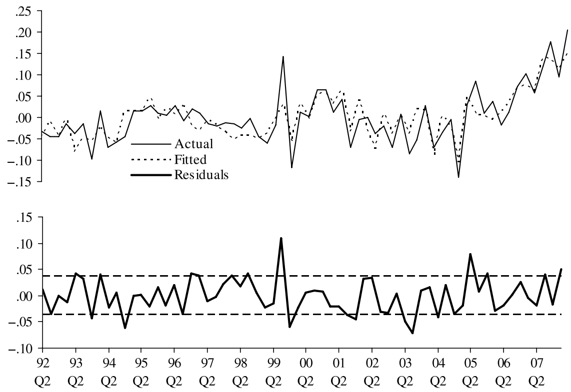

Fig. 3. Changes in log default rates for residential loans.

The adjustment coefficients suggest that just 10.6% of the deviation of the delinquency rate from its long-run path is corrected for in a quarter. A comparison with Table 5 shows this to be lower than for credit card delinquency or for other consumer loans. This is consistent with homeowners trying to maintain credit card repayments rather than real estate repayments in the short term if the short term equilibrium default rate increases to be above the long run equilibrium rate. One explanation for this is that payments which are missed on a mortgage are likely to be much larger than those for credit cards, and so the former are less easy to restore to their scheduled level from a given income. Another explanation is that there are readily available substitutes for buying a home, for example renting, but fewer substitutes for credit cards for many types of expenditures, although this may involve losing equity.

Fig. 3 shows the observed and predicted changes in default rates for residential loans. Clearly, the model underestimates the size of the increase in 1999 Q3 and overpredicts in the next quarter, and predicts a rise in 2005 Q1, which is one quarter early. It also underpredicts in 2005 Q3 and overpredicts in 2003 Q3. Note that the model fits the data just as well after 2004 Q4, when the default rate began to rise rapidly, as before.

7. Forecasting performance

In this section we examine the effects of shocks to the independent variables on the volume of delinquent consumer debt, and compare the accuracy of forecasts derived from the short-run dynamic model with those given by benchmark ARIMA models. We consider only the volume of delinquency consumer debt, because this is the only type of debt for which the volume of delinquent debt could be calculated.

Table 6
Short run dynamic mortgage equation.

Dependent variable:	Log Δ in mortgage delinquency rate (dlnrnsa)		
Estimation period:	1992(2)–2008(1)		
Independent variables		Coefficient	t-stat
Constant		0.042408	5.360**
Δpersonal loan interest rate	dlninsa(−2)	0.646873	2.289*
Δmortgage interest rate	dlnmisa	0.516384	4.831**
	dlnmisa(−1)	−0.432677	−4.569**
	dlnmisa(−2)	−0.310225	−3.202**
	dlnmisa(−3)	−0.390666	−3.665**
Δreal personal disposable income	dlnrpdisa	−2.401499	−4.128**
Δreal house price index	dlnrhpsa(−1)	−3.403896	−4.761**
Error correction	ecmlnrnsa12v1(−1)	−0.106041	−7.252**
	ecmlnrnsa12v2(−1)	0.400010	2.312*
Adjusted R^2		0.662679	
DW		2.232973	
Jarque-Bera $\chi^2(2)$		2.76793	
RESET2 $\chi^2(1)$		1.77440	
LM het. test $\chi^2(1)$		0.002740	

All variable changes are in logs.
* Significant at the 5% level.
** Significant at the 1% level.

Table 7
Summary statistics (1987 Q1–2008 Q1) for variables and shocks.

	Delinquency ($ million) (rdelsa)	Interest rate % (insa)	Credit (rccoutsa) ($ billion)	Sentiment index (sent)	Unemployment % (uetsa)	House price index (hpisa)
Minimum	12 845	11.59	411 750	110.37	3.93	142.11
Maximum	23 689	15.70	670 234	139.04	7.61	384.84
Average	17 823	13.53	524 853	127.23	5.46	230.25
Std. dev.	2 182	1.10	68 808	6.34	0.91	74.27
Shock value		1.00	70 000	15.00	1.00	75.00
Delinquency impact:						
Initial		321	4417	−331	0	0
Long-run		4983	6473	507	0	0
Most extreme		4987	6621	−589	442	−3805

Table 7 describes the values of the variables and indicates the shock to be applied to each in turn. The shock is typically set to roughly one standard deviation of the variable. The simulation starts from a situation where all of the variables are constant at their average values, and the trend variable is constant at the value that makes these average values consistent with long-run equilibrium. Each variable in turn is raised by the amount of the shock and held at that higher value indefinitely. Fig. 4 plots the cumulative response over 24 quarters.

All four of the independent variables are near their long-run cumulative impact within 12 quarters. Most do not move monotonically, perhaps reflecting some disorientation and transitional financial adjustments following the shock. The impacts shown in Fig. 4 shed further light on the plausibility of explanations of changes in delinquency. The impact of an increased outstanding debt is consistent with the credit quality explanation of the aggregate delinquency. At the level of household explanations, the adverse income shock explanation is consistent with the immediate and persistent increase in the delinquency volume when interest rates rise. It is also consistent with the impact of a shock to the unemployment rate, which leads to no changes in the delinquency volume over the first three months, perhaps whilst households are dissaving to fund repayments, but a dramatic increase in delinquency thereafter. The irrationality hypothesis is supported by the impact of a shock to optimism, which results in an ever increased delinquency after a period of three months, perhaps as households irrationally take on more debt. The strategic default hypothesis is supported by the impact of a positive shock to house prices that results in an immediate fall in delinquency volume, followed by a return to the original level as households adjust to the new levels.

The experience of credit repayment behaviour reflects the joint impact of ongoing shocks to all independent variables, and each shock response will occur before the responses to previous shocks have been exhausted. Fig. 2 demonstrates the model's success in coordinating these influences in order to track delinquency developments well, and in so doing gives credibility to the predicted responses to individual variables. The comprehensiveness of the model in doing so is indicated not only by the modest magnitude of its tracking errors, but by the absence of a evident pattern in these errors.

Fig. 5 compares the in-sample tracking properties of the regression model with those of an ARIMA model, in addition to the out-of-sample forecasts from both

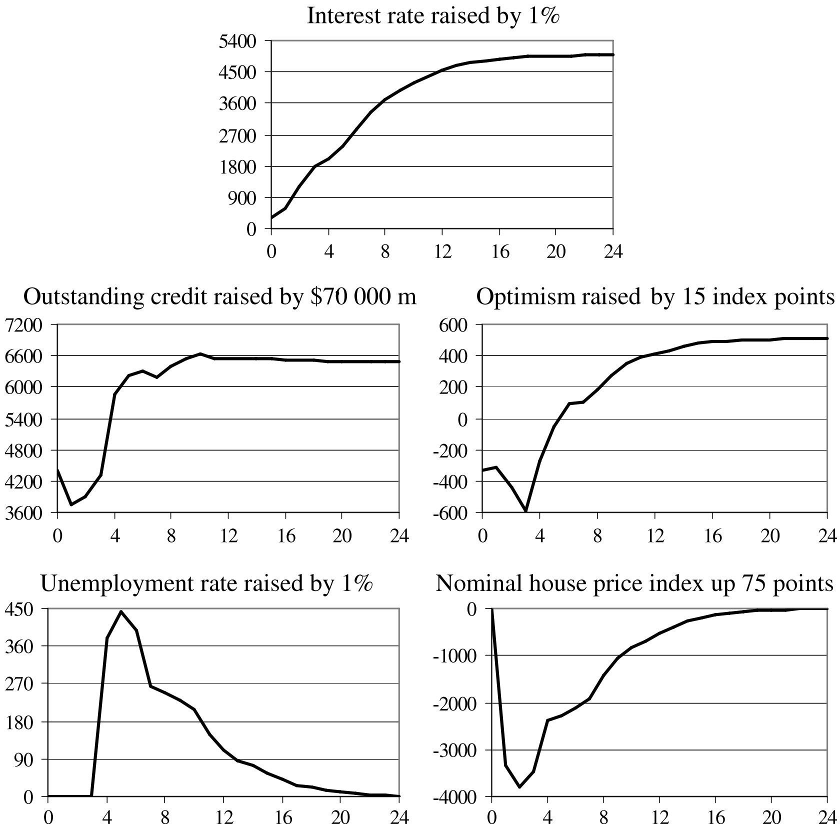

Fig. 4. Cumulative impact ($ million) on the volume of credit delinquency as a result of various shocks to the independent variables over 24 quarters.

models. Regression models are often valued for their analytical facilities, in spite of forecasting performances which are inferior to those of simple models that have little explanatory content, but which manage to extrapolate the trends and cycles in a dependent variable's behaviour well. Regression models are handicapped by the need to use forecasts of the independent variables in any out-of-sample predictions, and thus depend on the forecast errors of the independent variables being small or cancelling out. In order to assess the extent of this handicap, our model estimation reserved a holdout sample of observations for 2008 Q2–2009 Q1. In this four-quarter period, the model will have access to actual observations only to the extent that it is fitting lagged variables observed before the holdout period. Current observations and ones with short lags eventually require us to resort to forecast values. Table 8 shows the ARIMA models used to forecast the independent variables. The ARIMA models for the dependent variable establish a benchmark performance against which the regression model can be assessed.

In general, the ARIMA models reported in Table 8 reflect a suitable level of parsimony with respect to the numbers of estimated coefficients; however, marginally insignificant parameters are adopted occasionally as well, in order to achieve a suitably impressive ACF. To the extent that missed parsimony causes suboptimal forecasts of the independent variables, regression model forecasts will tend to appear in a poorer light than the benchmark forecasts.

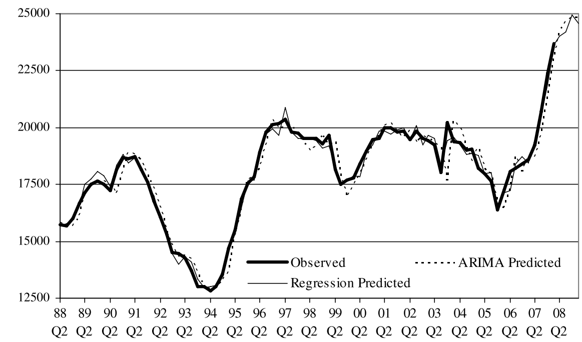

Fig. 5. Tracking alternative models for the delinquency volume.

Table 9 reports the out-of-sample forecast performances of the regression model and its two benchmark competitors. These are m-step-ahead forecasts that make no use of data observed in the out-of-sample period. Wherever an out-of-sample observation of a forecast is needed, relevant forecasts are used. For lagged independent variables in regression forecasts, the regression forecasts are used, and the relevant ARIMA forecasts are used for the other independent variables.

Not surprisingly, the model's out-of-sample performance has failed to achieve anything like the promise indicated by its in-sample performance. However, the delinquency volume regression model roughly matches the

Table 8
ARMA models for the first differences of log-transformed variables.

	Benchmark model		Forecasting models for predictor variables									
	Delinquency		Interest rate Δ [ln(insa)]		Credit Δ [ln(rccoutsa)]		Sentiment Δ [ln(sentsa)]		Unemployment Δ [ln(uetsa)]		House prices Δ [ln(hpisa)]	
Performance:												
R^2	0.319977		0.188680		0.522937		0.257891		0.437084		0.585216	
Std error	0.031070		0.017178		0.011287		0.027015		0.029082		0.005245	
Estimates:	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat
Constant:			−0.00278	−2.884**	0.00497	2.317*					0.01089	3.863**
AR1:	−0.24360	−2.247*							0.86152	8.130**		
AR2:	−0.30925	−2.711**			0.28798	2.504*						
AR3:					0.33734	3.167**					0.61883	5.849**
MA1:					−0.28052	−2.388*	0.24016	2.312*	0.51282	3.299**	−0.62447	−5.869**
MA2:			−0.20841	−1.744							−0.37341	−3.070**
MA5:					−0.24981	−1.950						
MA11:							−0.26818	−2.503*				
MA14:									0.23904	2.246*		
MA18:					0.36622	2.466*						
MA20:							−0.31990	−2.635*				
SAR1:	0.47243	2.159*			−0.27847	−2.481*						
SAR2:	−0.28147	−2.153*			−0.25491	2.191*						
SAR3:					−0.34215	−2.809**						
SMA1:	0.63859	2.946**	0.31898	2.655**			0.38022	3.526**	0.22641	1.844		
SMA1:			0.30804	2.449*					0.35156	2.952**		
Box–Liung prob:												
At lag24	0.606536		0.944222		0.936595		0.985520		0.969858		0.935956	
Min by lag24	0.493075		0.265556		0.671857		0.718427		0.870855		0.850769	

Note that the constants cited above are the non-zero estimated mean values for the series, not the intercepts.

Table 9
Comparison of regression forecasts with ARIMA benchmarks.

	"Actual" values	ARIMA forecast	ARIMA errors	Regression forecast	Regression errors
Ex-sample forecasts					
2008 Q2	**24 264**	**24 094**	171	**23 999**	266
2008 Q3	**25 111**	**24 675**	436	**24 216**	895
2008 Q4	**29 974**	**24 865**	5109	**24 960**	5014
2009 Q1	**33 733**	**24 832**	8901	**24 568**	9165
Out-of-sample					
RMSE			5137		5244
In-sample RMSE			562		268

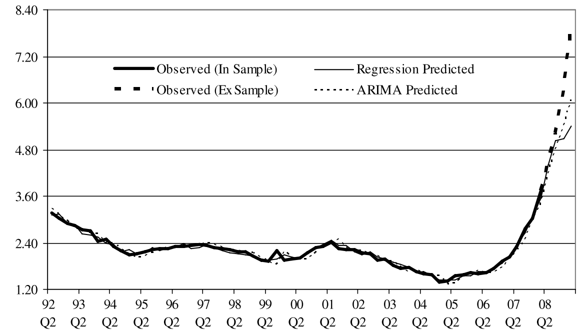

Fig. 6. Tracking alternative models for the real estate loan delinquency rate.

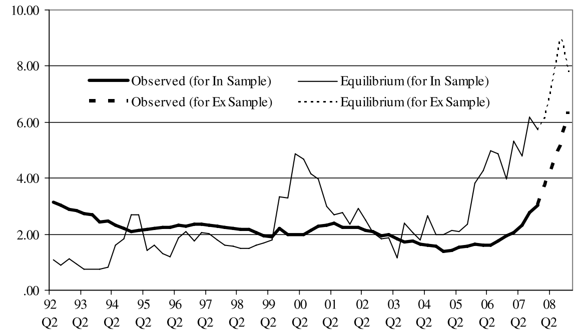

Fig. 8. Observed and equilibrium real estate loan delinquency rates.

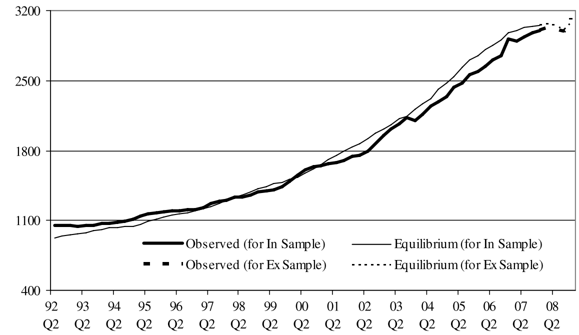

Fig. 7. Observed and equilibrium real estate loans outstanding ($ million).

out-of-sample performance achieved by the corresponding ARIMA model.

The present model explaining real estate loan delinquency rates is particularly interesting, because toxic real estate credit is at the heart of the current global financial crisis. Fig. 6 demonstrates that this model forecasts such delinquency quite well, even into the holdout period, following its considerable acceleration well into that period.

One might expect such a high level of fidelity of the model performance to reflect a simple appreciation of the extent to which real estate credit has been radically overextended in recent years, but in fact, that is not actually so evident. There is a large and significant coefficient on the error correction vector for the outstanding real estate debt, indicating that delinquency will be profoundly influenced by such debts being extended beyond equilibrium levels. However, Fig. 7 suggests that the mechanism is perhaps not quite so simple. The figure does show debt levels being persistently above equilibrium levels in recent years, but the relative magnitude of the excess seems modest and stable.

That the delinquency crisis reflects home purchases by people who are borrowing beyond their means can hardly be doubted, but the model suggests that the influence of income is a short-term dynamic phenomenon. There is no income variable in the cointegrating vector concerning the mortgage debt outstanding, and this reflects an absence of a long-run income relationship beyond what can be accounted for by proxy variables such as the trend and house prices. The more influential feature of the model is the small yet very significant coefficient on the error correction vector for real estate loan delinquency itself. Fig. 8 indicates that the delinquency rate is well out of equilibrium, and the model suggests that it will continue to grope upward for it for a while to come, unless there are shocks in the intervening time.

8. Conclusion

We have found evidence of a long-run relationship between the volume of delinquent consumer credit and the volume of consumer debt outstanding and the interest rate on personal loans. We have also found long-run relationships between delinquency rates for credit cards, the credit card interest rate and the level of household debt, as well as between delinquency rates for other debts and an index of household optimism. We also found a relationship between default rates on residential real estate loans, the mortgage rate and the real house price index.

These findings suggest that different explanations of delinquency are appropriate for different types of debts. For the volume of consumer debt, variations in the quality of debt, but not changes in the stigma of default, appear to drive the delinquent volume; while at the level of the household irrationality, adverse income shocks and changes in house prices are at work. Decomposing consumer debt into the debt on credit cards and that on other consumer loans, we find evidence that the quality of debt and averse shocks apply, but negative equity does not, whereas for other consumer loans it is only strategic delinquency that applies, which is plausible because credit card loans do not involve collateral whilst other loans often do. For real estate loans we find that both explanations of delinquency apply, as do all household level explanations.

These results are not consistent with the findings of Gross and Souleles (2002), who found evidence of reduced stigma in the case of credit cards. Our results are only partly consistent with those of Grieb et al. (2001); whilst we find evidence of the adverse shock hypothesis, we do not find that high interest rates, higher debt and unemployment significantly affect delinquency.

We also found that the error correction model gave forecasts of the volume of delinquent debt which were of an accuracy comparable to that of an ARIMA model.

Acknowledgement

We would like to thank Paul Mizen for helpful comments.

Appendix. Definitions of variables

Lnrdelsa Log of (real consumer loan debt outstanding on loans to US chartered commercial banks which is 30+ days overdue, in $00 millions at year 2000 prices).
Sources of raw data: *Charge off and delinquency rates on loans and leases at commercial banks*, Consumer loans: All and Series G19. *Consumer Credit* debt outstanding to commercial banks. All series are from FRB.
Seasonally adjusted by the authors using X12.

Lnccsa Log of (consumer credit card debt to US chartered commercial banks which is 30+ days overdue, as a percentage of the end-of-period corresponding debt outstanding).
Sources of raw data: *Charge off and delinquency rates on loans and leases at commercial banks*, Consumer loans: Credit cards, FRB.
Seasonally adjusted by the authors using X12.

Lnosa Log of (consumer non-credit card debt to US chartered commercial banks which is 30+ days overdue, as a percentage of the end-of-period corresponding debt outstanding).
Sources of raw data: *Charge off and delinquency rates on loans and leases at commercial banks*, Consumer loans: Other, FRB.
Seasonally adjusted by the authors using X12.

Lnrnsa Log of (single family residential mortgage debt (including home equity loans) to US chartered commercial banks which is 30+ days overdue, as a percentage of the end-of-period corresponding debt outstanding).
Sources of raw data: *Charge off and delinquency rates on loans and leases at commercial banks*, Real Estate Loans: Residential, FRB.
Seasonally adjusted by authors using X12.

Lninsa Log of (nominal interest rate on 24 month personal loans).
Source of raw data: *Terms of Credit, Consumer Credit Historical Data*, FRB.
Seasonally adjusted by the authors using X12.

Lnrccoutsa Log of (sum of the revolving and non-revolving consumer credit outstanding to commercial banks in $00 millions, divided by the seasonally adjusted price index personal consumption expenditure (2000 = 100)).
Sources of raw data: FRB *Historical Consumer Credit Data, Major Types of Credit* and Bureau of Economic Analysis, *Price Indices for Personal Consumption Expenditures by Major Type of Product*, Table 2.3.4.
The numerator was seasonally adjusted by the authors using X12.

Lnrpdisa Log of (disposable personal income (in $00 millions), seasonally adjusted, divided by the price index of the personal consumption expenditure, seasonally adjusted (2000 = 100)).
Sources of raw data: *Price Indices for Personal Consumption Expenditures by Major Type of Product*, Table 2.3.4, and *Personal Income and its Disposition*, Table 2.1, Bureau of Economic Analysis.

Lnsent Log of the index of the relative expected change in financial situation in one year's time, relative sentiment.
Source: Index of Consumer Sentiment, Table 6 Expected Change in Financial Situation, *Index of Sentiment, Surveys of Consumers*, Institute for Social Research, University of Michigan.
Seasonally adjusted by the authors using X12.

Lnhpisa Log of (US combined house price index seasonally adjusted).
Sources of raw data: *OFHEO House price index, US Combined Index*: Office of Federal Housing Enterprise Oversight Office.
The OFHEO house price index was seasonally adjusted by the authors using X12.

Lnrhpisa Log of (US combined house price index, seasonally adjusted, divided by the price index personal consumption expenditure, seasonally adjusted (2000 = 100)).
Sources of raw data: *OFHEO House price index, US Combined Index*: Office of Federal Housing Enterprise Oversight Office*; Price Indices for Personal Consumption Expenditures by Major Type of Product*, Table 2.3.4, Bureau of Economic Analysis.
The OFHEO house price index was seasonally adjusted by the authors using X12.

Lnrnoutsa Log of (real estate loans outstanding to commercial banks/price index).

Source: Series bcablcr_ba.m, *Federal Reserve Board*. The numerator was seasonally adjusted by the authors using X12.

Lnmisa Log of (nominal interest rate on conventional conforming 30-year fixed rate mortgages). Source: *Primary Mortgage Market Survey*, Freddie Mac. Seasonally adjusted by the authors using X12.

Lnccinsa Log of (nominal credit card interest rate). Source: Consumer Credit G19, Terms of Credit, *Federal Reserve Board*. Seasonally adjusted by the authors using X12.

Lndsrsa Log of (debt service ratio). (Ratio of household debt payments to disposable personal income.) Source: *Federal Reserve Board*. Seasonally adjusted by FRB.

Lnrtdoutsa Log of (total credit market debt owed by the household sector, seasonally adjusted, divided by the price index of the personal consumption expenditure, seasonally adjusted (2000 = 100)). Source: *Federal Reserve Board*, Flow of Funds Accounts of the United States, Outstandings, file ltab1d.prn, series FL154102005.Q and Bureau of Economic Analysis, *Price Indices for Personal Consumption Expenditures by Major Type of Product*, Table 2.3.4. The numerator was seasonally adjusted by the authors using X12.

All seasonal adjustments were performed before logs were taken.

References

Agarwal, S., & Liu, C. (2003). Determinants of credit card delinquency and bankruptcy: macroeconomic factors. *Journal of Economics and Finance*, *27*(1), 75–84.

Arner, D. W. (2009). *The global credit crisis of 2008: the causes and consequences*. AIIFL working paper no. 3.

Bank of International Settlements (2006). *International convergence of capital measurement and capital standards: a revised comprehensive framework*. Basel: Bank of International Settlements.

Bellotti, T., & Crook, J. N. (2009). Credit scoring with macroeconomic variables using survival analysis. *Journal of the Operations Research Society*, *60*(12), 1699–1707.

Buck, B. K., Kinneckel, A. B., & Moore, K. B. (2006). Recent changes in US family finances: evidence from the 2001 and 2004 survey of consumer finance. *Federal Reserve Board Bulletin*.

Crouhy, M. G., Jarrow, R. A., & Turnbull, S. M. (2008). *The subprime credit crisis of '07*. Available at SSRN: http://ssrn.com/abstract=1112467 (09.07.08).

Deaton, A. (1992). *Understanding consumption*. Oxford: Clarendon Press.

Deng, Y., Quigley, J. M., & Van Order, R. (2000). Mortgage terminations, heterogeneity and the exercise of mortgage options. *Econometrica*, *68*(1), 275–307.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error corrections representation, estimation and testing. *Econometrica*, *55*(2), 251–276.

Figuera, C., Glen, J., & Nellis, J. (2005). A dynamic analysis of mortgage arrears in the UK housing market. *Urban Studies*, *42*(10), 1755–1769.

Grieb, T., Hegji, C., & Jones, S. T. (2001). Macroeconomic factors, consumer behaviour, and bankcard default rates. *Journal of Economics and Finance*, *25*(3), 316–327.

Gross, D., & Souleles, N. (2002). An empirical analysis of personal bankruptcy and delinquency. *Journal of Financial Studies*, *15*(1), 319–347.

Heitfield, E. (2005). Dynamics of ratings systems. In *Basel committee on banking supervision, studies on the validation of internal ratings systems*. Working paper 14. Basel: Bank of International Settlements.

Johansen, S. (1988). Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254.

Kau, J. B., Keenan, D. C., & Kim, T. (1993). Transaction costs, suboptimal termination and default probabilities. *Journal of the American Real Estate and Urban Economics Association*, *21*(3), 247–263.

Kau, J. B., Keenan, D. C., & Kim, T. (1994). Default probabilities for mortgages. *Journal of Urban Economics*, *35*(3), 278–296.

Lambrecht, B., Perraudin, W., & Satchell, S. (1997). Time to default in the UK mortgage market. *Economic Modelling*, *14*, 485–499.

Laibson, D., Repetto, A., & Tobacman, J. (2003). A debt puzzle. In P. Aghion, R. Frydman, J. Stiglitz, & M. Woodford (Eds.), *Knowledge, information, and expectations in modern economics: in honor of Edmund S. Phelps*. Princeton: Princeton University Press.

Patterson, K. (2000). *An introduction to applied econometrics: a time series approach*. New York: St. Martin's Press.

Stiglitz, J. E., & Weiss, A. A. (1981). Credit rationing in markets with imperfect information. *American Economic Review*, *71*, 393–410.

Sullivan, A. C. (1987). *Economic factors associated with delinquency rates on consumer instalment debt. Credit research center*. Working paper no. 55. Krannert Graduate School of Management, Purdue University.

Teo, A. H. L. (2004). Delinquency risk in residential ARMs: a hazard function approach. *Journal of Real Estate Portfolio Management*, *10*(3), 243–258.

Whitley, J., Windram, R., & Cox, P. (2004). *An empirical model of household arrears*. Working paper 214. Bank of England.

Jonathan Crook is the Director of the Credit Research Centre at the University of Edinburgh (http://www.crc.man.ed.ac.uk) and Professor of Business Economics. He is the joint author/editor of several books, including: *Readings in Credit Scoring*, OUP 2004, and *Credit Scoring and its Applications*, SIAM 2002. He has published a large number of papers relating to credit in many different academic journals. He is a Fellow of the Wharton Financial Institutions Center, Wharton School, University of Pennsylvania, and has been a postdoctoral Fulbright Scholar at the University of Virginia and a Visiting Fellow at other universities.

John Banasik is a Senior Lecturer (Associate Professor) in Management Science and Statistics in the Business School, University of Edinburgh. He has published extensively on historical time series econometric models and credit scoring models, especially reject inference. In 1999 he was joint winner (with Crook and Thomas) of the OR Society's Goodeve Medal.