Punishment-Induced Deterrence: Evidence from the Video-Rental Market

Peter Fishman
Department of Economics
University of California, Berkeley

Devin G. Pope
Department of Economics
University of California, Berkeley

This Draft: May 26, 2006
(Please do not cite without authors' permission)

Abstract

Does the memory of recently being punished deter criminals from committing crimes? Criminologists have long discussed the psychological effect that receiving a punishment can have on future criminal behavior. While it may exist anecdotally, this psychological deterrent effect is difficult to disentangle from classical deterrence in a real-world setting because of changes in information and incentives that typically occur when an individual is punished. In this paper, we test for punishment-induced deterrence in a controlled market where issues of changes in expected benefits and costs can be addressed: the video-rental market. We explore the effect of having to pay a late fee on customer behavior and find evidence of negative state dependence. Specifically, we find that paying a late fee reduces the probability of paying a late fee in the subsequent visit by 19% and that this deterrent effect decays quickly over time. We show that this behavior is not mitigated by experience and discuss the implications of these findings on consumer rationality, optimal crime policy, and marketing.

*We thank David Card, Ken Chay, Stefano DellaVigna, Robert Fishman, Kory Kroft, David Lee, Jim Powell, Matthew Rabin, and seminar participants at UC Berkeley for helpful comments and suggestions. We are grateful to the video-store owner that provided the data and the IBER for funding. All errors are our own.

I. Introduction

Economists have long studied how individual behavior is shaped by incentives. It is typically assumed that decisions are made by forward-looking, rational agents who consider the costs and benefits associated with all possible actions. This framework is exemplified in Becker's (1968) seminal paper on crime. Becker argued that criminal action is driven by cost-benefit analyses, suggesting that a policymaker can deter criminal behavior by creating a high "price" of crime. In Becker's model, the price of crime is determined by the probability of detection along with the punishment imposed if detected. The ability to deter individuals from committing crimes by increasing the expected punishment is what we refer to as incentive-induced or classical deterrence.

In this analysis, we focus on an additional mechanism of deterring criminal behavior which we label *punishment-induced deterrence*. We consider punishment-induced deterrence to be the subsequent deterrent effect on behavior that actually experiencing punishment for a crime has on the specific individual who was punished, *conditioning for changes in expected benefits and costs of future criminal activity*.[1] Early followers of utilitarianism claimed that punishment deterred criminal behavior because people responded to a subjective pleasure/pain calculus (Bentham, [1843] 1962). Glueck (1928) noted the distinction between deterrence effects from the fear of future punishment (classical deterrence) and the memory of past punishments (punishment-induced deterrence). Punishment-induced deterrence does not fit into a model of

[1] Punishment-induced deterrence is similar to what has been labeled as specific deterrence in the sociology and criminology literatures. Specific deterrence is typically defined as the deterrent effect of being punished on the specific individual. However, specific deterrence is considered the conglomeration of what we call punishment-induced deterrence and classical deterrence because of changes in information or incentives that often occur when an individual is punished. While we provide a discussion and literature review of specific deterrence in the next section, we consider punishment-induced deterrence to be a distinct topic.

forward-looking agents choosing whether or not to take an action based solely on the expected benefits and costs. Rather, it implies that agents are affected psychologically by experiencing the consequences of their actions that have occurred in the recent past causing them to give additional attention to avoiding the punishment in the future.

There are many situations in which a punishment-induced deterrent effect seems intuitive. Getting a traffic ticket may induce an individual to be a more careful driver for a period of time. A basketball player who fouls out in one game may be extra careful with fouls in the next few games. A shoplifter who is caught and punished may be less likely to steal in the future relative to a shoplifter who is not detected.

A fundamental problem with the identification of punishment-induced deterrence in these and other examples is separating its effect from that of classical deterrence. For example, there are three main ways in which experiencing a punishment can have a classical deterrent effect through the changing of future costs and benefits of committing crime. First, being detected and punished for committing a crime may provide new information to agents who can then update beliefs. For example, being caught speeding may provide a driver with information about the number of police patrolling the roads. Thus, driving more carefully after receiving a traffic ticket may be a very rational response to the updated costs and benefits. Second, being detected and punished for a crime oftentimes changes the punishment that will be received if the individual commits the same crime again. The cost of receiving a speeding ticket may be higher if it is the second ticket received by the individual (insurance rates may increase more for the second ticket, most states have laws that mandate a license suspension for anyone who receives a certain number of tickets in a certain amount of time, etc.). Third, being punished for serious crimes typically results in prison or jail time, making it impossible to repeat the offense for a period of time. Thus, because of an incapacitation effect, there is a mechanical relationship between committing a crime and subsequently not committing a crime for a period of time. Because of changes in information, differences in punishments for repeat offenders, and incapacitation, observing less crime from individuals following a punishment might be mistakenly considered punishment-induced deterrence when in reality it is simply mechanical or the effect of classical deterrence.

In order to circumvent these issues and obtain an accurate measure of punishment-induced deterrence, we abstract from actual criminal decisions by using data from the video-rental market. Analyzing a large, individual-level dataset of movie-rental and return decisions, we consider the act of returning a movie late to be a "crime" and having to pay a late fee a "punishment." Analyzing movie-return decisions allows us to overcome the common obstacles, discussed above, that make it difficult to disentangle punishment-induced from classical deterrence: receiving a video late fee (punishment) provides no information regarding the probability of being detected in the future since detection is constantly 100%; the punishment is known and remains unchanged regardless of previous offenses; and problems due to incapacitation do not exist. Thus, the video-rental market provides a unique environment where punishment-induced deterrence can be identified using market decisions.

Using a semiparametric econometric technique in order to control for unobserved individual-specific effects in the dynamic process, we test whether receiving a late fee affects the propensity to return videos late in subsequent periods. As predicted by punishment-induced deterrence, we find evidence that there exists negative state dependence in late-fee payments. Our results indicate that paying a late fee reduces the probability that the customer will pay a late fee in the subsequent visit by 19% (2.6% off a base rate of 14%). A smaller deterrent effect of about 8% is found from receiving a late fee on the probability of receiving a late fee two periods later suggesting that the effect of punishment-induced deterrence decays as time passes. We also provide evidence that having to pay a large late fee results in a greater deterrent effect than paying a small late fee. Looking at a different deterrence dimension, we test for the effect of paying a late fee on the number of days before a customer returns to rent more movies and the number of movies that are rented during the subsequent visit. While we find no evidence that individuals adjust the number of movies they rent after receiving a late fee, we do find that they delay returning to the video store by an additional 0.73 days. This result also decays to the point of statistical insignificance after two periods.

A framework is provided for how punishment-induced deterrence can be included into a model of criminal behavior and its implications on optimal crime policy are discussed. We discuss how the existence of punishment-induced deterrence implies higher optimal expected punishments relative to those found under the classical model of crime. While in practice, punishment-induced deterrence is often captured along with classical deterrence when studying the effect of a policy change, we discuss situations in which it may be ignored, resulting in inefficient punishment levels. Another important implication of punishment-induced deterrence is the way in which punishments should be meted out. Linking punishment-induced deterrence to the literature on attention, we argue that making punishments more salient will lead to a larger deterrence effect. This claim has implications for not only policymakers, but also to marketers at firms whose aim may be to minimize punishment-induced deterrent effects (e.g. video stores). The firm, which wants to deter as few of its customers' "crimes" as possible, should make punishments less salient.

Our results have additional implications for understanding consumer behavior. Punishment-induced deterrence implies that outcomes that occurred in the recent past are able to affect an individual's behavior even after controlling for changes to incentives that the past outcomes could have caused. If behavior changes as a result of an individual being in a state of having recently been punished, individuals are making a "mistake" either before or after the punishment.[2,3] It has been shown in many instances that experience causes individuals to converge to rational behavior (List, 2003, List, 2004). However, our results indicate that punishment-induced deterrence is not mitigated by experience. Even customers who have rented videos and paid late fees numerous times exhibit changes in behavior every time a salient reminder of the consequences of returning a movie late is provided in the prior period.

In the next section we discuss previous papers that have addressed the related topic of specific deterrence. We also discuss the attention literature and how it relates to the behavior being discussed in our analysis. Section III provides a conceptual

[2] Punishment-induced deterrence can also be explained by a model of learning and forgetting. For example, receiving a video late fee may provide information to an individual about himself or about how much the he dislikes the punishment meted out resulting in a decrease in the propensity to receive late fees in the future. This response may be considered rational and in accordance with classical deterrence. However, to consider this behavior rational, even after renting and paying late fees several times, a video-rental customer must still be learning and perhaps more importantly forgetting what was learned a few visits after receiving a late fee. While this model of learning and forgetting is observationally equivalent to the backward-looking behavior story that we discuss in this paper, the implications are similar and both stories imply a lack of sophistication by consumers.

[3] Another explanation for observing negative state dependence in video rentals is credit constraints. Getting a late fee may cause an individual to be more careful in a subsequent visit because discretionary spending is now limited. The credibility of this explanation is limited due to the small size of the fines in this application as well as the fact that the video store from which our data are drawn is located in a wealthy neighborhood in the Bay Area.

framework and discusses optimal crime policy. Section IV describes the data and our empirical strategy. The results are presented in Section V and Section VI provides a discussion and conclusion.

II. Related Literature

Several studies have estimated the behavioral response to changes in the expected punishment of committing a crime (Kessler and Levitt, 1999, Levitt, 1996, Levitt, 1997, Katz, Levitt, and Shustorovich, 2003, Di Tella and Schargrodsky, 2003 and Lee and McCrary, 2005). The results of these studies, which test for classical deterrence, typically find that individuals react to changes in the probability of detection or the punishment associated with committing a criminal act.[4]

Specific deterrence has received attention in both the sociology and criminology literature. This literature defines general deterrence to be the deterrent effect that all individuals face because of expected detection rates and punishments. Specific deterrence, on the other hand, is referred to as the subsequent deterrent effect that a specific individual may exhibit after being punished for a crime. Many of the papers that study specific deterrence have in mind that being punished provides a reality check or in some other way psychologically deters people from committing additional criminal acts. However, the literature does not typically attempt to distinguish this psychological deterrent effect from behavior changes due to changes in information, incentives, or incapacitation that may simultaneously occur with being punished.

[4] A notable exception is Lee and McCrary (2005) who find no change in crime rates when the punishment exogenously changes.

A common approach to test for specific deterrence is to compare the number of repeat offenses committed by individuals who have had police contact (but were not arrested) to individuals who were arrested (for similar offenses). These studies have produced mixed results. Clarke (1966), Cohen and Stark (1974), McCord (1983), Sherman and Berk (1984), and Smith and Gartin (1989) all find evidence suggestive of specific deterrence while Gold and Williams (1969), Shoham (1974), Farrington (1977), and Klemke (1978) find no evidence (or evidence in the opposite direction) of specific deterrence. While these studies can provide insight into the effects of different crime policies, they are unable to disentangle the effect of what we call in this paper punishment-induced deterrence from that of classical deterrence due to the changes in information, repeat-offense punishments, or incapacitation that may occur between the treatment and control groups.

In the economics literature, Chen and Shapiro (2005) use a regression-discontinuity design to identify the effect of being placed in a harsher prison environment on recidivism. They find that individuals who are placed in rougher prison environments are more likely to commit a crime after being released than their counterparts (evidence in the opposite direction of specific deterrence). Pintoff (2005) also uses a regression-discontinuity design in order to identify the effect of incarceration on juvenile recidivism. She finds that exogenous incarceration causes a large reduction in the probability that individuals will repeat offend (consistent with specific deterrence) even when controlling for incapacitation effects. However, whether the effects that these papers find are due to punishment-induced deterrence, as defined in this paper, or rather, simply responding to the updated benefits and costs of repeat offending is unclear.

The literature on limited attention can provide insight into the underlying psychology behind punishment-induced deterrent effects. Each day, people make a myriad of decisions each with expected consequences. Given limited attention, it has been argued that individuals will pay attention to expected consequences that are in some way more salient than others (Fiske and Taylor, 1991). Thus, the basic prediction of the theory of limited attention is that agents will pay too much attention to salient stimuli (Barber and Odean, 2004 and Huberman and Regev, 2001) and too little attention to non-salient stimuli (Hong, Torous, and Valkanov, 2002, DellaVigna and Pollet, 2006). In some ways, "attention-induced deterrence" could be an equally appropriate term to use in this paper. It is possible that simply focusing an individual's attention (even without changing information or incentives) to the negative consequences of one's actions can influence an individual's choice. Perhaps the most obvious way to shift someone's attention to focus on the potential punishments of committing a crime is actually punishing them when they do commit that crime. Our analysis uses this approach by analyzing how the saliency involved with experiencing a punishment (such as paying a fine) can cause individuals to devote additional attention to avoiding this punishment in the future, consistent with the prediction of punishment-induced deterrence. Two additional predictions of this attention story are that the larger the fine, the more attention-grabbing it will be, and as time passes and the stimuli caused by experiencing a punishment becomes less salient, the effect of punishment-induced deterrence will fade.

III. Conceptual Framework

Consider an agent who in time t faces the decision of whether or not to commit a crime. The following notation is used.

U_t = utility of wealth function in period t.

b_t = benefit of committing the crime in period t

p_t = probability of detection in period t

f_t = fine (punishment) collected from individual *if* detected in period t

F_t = fine (punishment) collected from individual in period t, $F_t = \{0, f_t\}$

y_t = initial income in period t

According to the classical theory of crime (Becker, 1968), the agent will commit the crime if the expected utility of doing so exceeds the utility of the endowed income. Thus, the agent will engage in the criminal activity if

(1) $(1-p_t)U_t(y_t+b_t)+p_tU_t(y_t+b_t-f_t)>U_t(y_t)$.

Simplifying this problem by assuming linear utility, the agent will commit the crime if

(2) $b_t - p_t f_t > 0$.

Alternatively stated, the criminal will commit the crime if the expected value is positive.

We now consider including punishment-induced deterrence in the criminal's decision problem. Punishment-induced deterrence predicts that even when using the individual's updated beliefs regarding b_t, f_t, and p_t, fines that were received in the past will have an effect on the current decision of whether or not to commit a crime. We model this by defining an "internalized punishment" as the transformed expected fine an agent considers when deciding whether to commit a crime. The punishment may be more or less salient and thus be considered bigger or smaller depending on whether or not the individual was recently punished. A higher internalized punishment is indicative of the punishment looming large in the individual's mind.[5]

(3) $\text{Internalized Punishment}_t = p_t \cdot f_t \cdot g_t(F_{t-1}, F_{t-2}, \ldots)$

The classical model of crime is the special case where $g_t(\bullet) = 1$ for all possible values of $F_{t-1}, F_{t-2}, \ldots$.

The existence of punishment-induced deterrence implies that

(4) $g_t(F_{t-1}, F_{t-2}, \ldots) > g_t(0,0,\ldots)$ if $F_{t-x} > 0$ for some $x \geq 1$.

Thus, receiving a fine in some previous period causes the internalized punishment in period t to be larger than it otherwise would have been. Drawing from the literature on attention, the functional form of $g_t(\bullet)$ can be conjectured to satisfy the following

(5) $$\frac{\partial g_t(\bullet)}{\partial F_{t-1}} > \frac{\partial g_t(\bullet)}{\partial F_{t-2}} > \ldots > 0.$$

This implies that punishment-induced deterrence is positively correlated with the temporal proximity of past punishments and that larger past fines result in a larger punishment-induced deterrent effect than smaller past fines.

Equation (2) can now be generalized to include the case of punishment-induced deterrence. The risk-neutral agent will engage in criminal activity if

(6) $b_t - p_t f_t g_t(F_{t-1}, F_{t-2}, \ldots) > 0$

[5] The internalized punishment as defined allows for individuals that recently received a punishment to give a higher weight to either the probability of detection or the fine. However, in the example used in this paper, video late fees, the probability of detection is 100%. Thus, for our application, it is more reasonable to think that the fine is looming large rather than the probability of detection looming higher than 100%.

Equation (4) and (6) imply that detecting and giving a fine to an individual who committed a crime in a previous period decreases the chance that the individual will commit a crime in the current period.

The function $g_t(\bullet)$ can be easily included in models such as those by Polinsky and Shavell (1979, 1991) to calculate optimal fine and detection rates. Under punishment-induced deterrence, there is a higher benefit of detecting and punishing an individual for committing a crime since the punishment results in a costless deterrent effect in the future. This increase in the marginal benefit of punishment (while holding the marginal cost of punishment constant) implies an optimal punishment level that increases with the magnitude of punishment-induced deterrence.[6] In this way, punishment-induced deterrence can be thought of as being similar to a peer effect.[7] Being punished for committing a crime decreases the probability that your future selves (peers) will commit a crime. Punishment-induced deterrence implies a peer-group-type effect that works within a single person across time rather than within peer groups across space.

The discussion of optimal fines and detection rates is also relevant to marketers. Many firms such as video stores, credit-card companies, and banks rely on fees generated by customers' delinquent behavior as an important source of revenue. Unlike a policymaker, the objective of these firms may be to increase "crimes" committed by their customers. Punishment-induced deterrence implies that giving a fine to an individual will cause the individual's future selves to be more careful. Thus, relative to a model

[6] This result holds for a policymaker whose sole objective is to set the marginal benefit from reducing crime equal to the marginal cost of detecting and punishing individuals.
[7] See Sah (1991), Case and Katz (1991), Glaeser, Sacerdote, and Scheinkman (1996), and Katz, Kling, and Liebman (2001) for a discussion of peer effects and crime.

limited to classical deterrent effects, a firm may choose to set a smaller fine (or detection rate if possible).

A policymaker assessing the costs and benefits of detection must properly account for punishment-induced deterrent effects. In the presence of both punishment-induced and classical deterrence, money spent on increased detection can result in two separate benefits: firms will decrease pollution immediately due to the forward-looking behavior of classical deterrence; in addition, firms that were detected and punished will subsequently reduce pollution levels due to the backward-looking behavior of punishment-induced deterrence. As long as a policy study uses a window of time that extends long enough past the policy change, punishment-induced deterrent effects will be included in the overall deterrent estimate. However, imagine the situation where the effect of the increased detection rates was measured by looking at only one pollution observation per firm after the new detection policy was put into place. The researcher would be missing the subsequent punishment-induced deterrent effects that the policy had on the punished firms. If these effects are not captured by the researcher, it may be concluded that the benefit of a higher detection rate is not worth the cost and thus the optimal expected punishment level would be set too low.

The attention story to which we attribute punishment-induced deterrent effects also suggests that policymakers may deter crime not only through adjusting punishment levels and detection rates, but by changing the saliency of the punishments. By making punishments more salient (even if the actual fine or time in jail does not change), individuals may further reduce future criminal acts. Jolls, Sunstein, and Thaler (1998), for example, discussed the deterrent value of parking tickets being delivered in bright green or orange envelopes with "VIOLATION" written in large letters as opposed to a more cost-effective envelope discretely placed on a car. Punishment-induced deterrent effects suggest that rather than a police officer giving a verbal warning to a juvenile who is caught vandalizing, the officer should make the punishment of being caught more salient (e.g. briefly handcuffing the offender). Giving punishments in this way might not change the future benefits and costs to the individuals involved, but the saliency of the event may provide a punishment-induced deterrent effect. Marketers on the other hand should implement policies that reduce the saliency involved with punishments (fees) in order to maximize profits. Automatic withdrawal or prepaid late-fee accounts may reduce the amount of punishment-induced deterrent effects and hence, increase subsequent delinquent behavior and total revenue from such behavior.

One caveat to the optimal policy implications discussed above concerns the effect of punishment-induced deterrence on individual welfare. Does punishment-induced deterrence cause individuals to make more or less efficient individual decisions? Consider an individual who has never been punished for committing a crime. If $g_t(0,0,...) < 1$, the criminal is myopic and is underestimating the true cost associated with committing a crime. By psychologically increasing the internalized punishment through detection and punishment, a policymaker is inducing the criminal to behave more efficiently. However, if the criminal is correctly weighing the cost of committing a crime, $g_t(0,0,...) = 1$, or already overweighing, $g_t(0,0,...) > 1$, then the added punishment-induced deterrent effect achieved through detecting and punishing will actually cause the criminal to make a less efficient individual decision.[8] Without knowledge of $g_t(0,0,...)$, we cannot speak to the effect of punishment-induced deterrence on individual welfare.

IV. Data and Empirical Strategy

We use a large dataset on video store transactions received from a large, independent video store in Northern California. The data set includes all transactions made by over 10,000 distinct customers during a two-year period from January 1st, 2003 through December 31st, 2004.[9] Each observation involves the set of transactions by an individual on a given day. For each observation, we have the account number, date, title of each movie rented, type of rental (new release, etc.), rental cost, the amount of money paid to cover a late fee for a past rental, and payment method (credit, cash, check, gift card). Using the account number, we are able to follow the rental behavior for a given individual over the two-year period.[10]

The video store for which we have data classifies movies into two categories: new and old releases. New releases have a one-day rental period while old releases are five-day rentals. Each additional day beyond the rental period for which a movie is not returned is associated with a late fee of \$3.00 for new releases and \$1.00 for old releases.[11] For each visit to the video store, we observe whether or not the customer paid money to cover a late fee associated with a previous rental (as opposed to observing

[8] For example, some people may think that people worry too much about small fines (e.g. paying a parking meter or buying a ticket to ride the metro when the expected value is negative).
[9] The first two observations for each individual were dropped in order to stagger the data in a way that the beginning of the year would not be the initial condition for a large fraction of renters.
[10] It is possible that multiple individuals share one account. We are unable to identify which accounts have multiple users. Sharing of accounts would most likely cause us to attenuate the true effect of punishment-induced deterrence since it induces noise in who is actually receiving the late fee (punishment).
[11] The video store allows customers to pre-purchase video rentals (at a slight quantity discount) and then use those pre-purchases to pay for late fees. In these cases (7.8% of the time), the late fee that the customers actually pay can be \$2.48 or \$2.75 instead of \$3.00.

which movies were returned late). The policy at this particular video store is that customers are asked to pay any late fees accrued from the previous rental whenever attempting to rent videos. If a customer returns a movie late and rents another movie in the same visit, they are asked at that time to pay the late fee. Thus, we associate paying a late fee in period t with movies returned late in period t-1. Occasionally, customers will return a movie late and decide to pay the late fee without renting any additional videos (2.6% of late fees are paid in this manner). Because they did not rent a movie when they paid the late fee, it will be impossible for them to have to pay a late fee during their subsequent visit. This behavior would mechanically provide evidence in favor of punishment-induced deterrence. To address this problem, we drop all observations which represent a visit to the video store in which a late fee was paid but no movie was rented.

Table 1 presents summary statistics for our data. As can be seen, the average person in our dataset rents 2.3 movies per visit and visits the video store 21 times during the two-year period. The movies are returned late 14% of the time causing the average individual to pay $16.50 in late fees over the two-year period.

We specify a simple model for late fee behavior,

(7) $$\text{Paid Fee}_{\text{it}} = \alpha_i + \gamma \, \text{Paid Fee}_{\text{it-1}} + \mu_{it},$$

where $\text{Paid Fee}_{\text{it}}$ is an indicator that equals one if individual i paid a late fee during video-store visit t, $\text{Paid Fee}_{\text{it-1}}$ is an indicator that equals one if individual i paid a late fee during her previous video-store visit (t-1), α_i is an unobserved individual-specific effect, and μ_{it} is a random disturbance that is i.i.d. over time. This model implies that after controlling for the type of each individual and last period's outcome, late fees are determined by transitory shocks. Punishment-induced deterrence represents a decrease in the probability of receiving a late fee in the current period due to the receipt of a late fee in the previous period. Thus, the hypothesis of punishment-induced deterrence implies that $\gamma < 0$.

We address two specific questions regarding the model specification. First, are fixed effects needed in this situation, especially considering the increased difficulties they cause in estimation? The video-store data used in this analysis suggests substantial customer heterogeneity in late-fee rates implying the existence of unobserved individual-specific effects. Column (1) of Appendix Table A presents the results from the simple regression of $\text{Paid Fee}_{\text{it}}$ on $\text{Paid Fee}_{\text{it-1}}$ using a linear probability model with no fixed effects. As would be expected if unobserved effects were an issue (and contrary to the hypothesis of punishment-induced deterrence), receiving a late fee during the previous visit increases the chance of paying a late fee during the current visit by 15.4%.

Second, we have assumed μ_{it} to be i.i.d. over time as opposed to allowing for serial correlation. Our intuition suggests that after controlling for unobserved individual-specific effects, serial correlation is a minor issue. However, one might imagine that if individuals have certain periods in their life that are particularly busy or relaxed (e.g. holidays), returning videos late may be positively correlated across time. If there is positive serial correlation in our data, we will be underestimating the effect of punishment-induced deterrence (negative state dependence). In order to overstate the case of punishment-induced deterrence, the less plausible story of negative serial correlation is required.

Econometricians have devoted much attention to the estimation of dynamic linear models with an additive unobserved effect. Ordinarily, a fixed effects framework would be ideal to control for a situation in which there exists individual heterogeneity. However, since a lagged dependent variable is used as an explanatory variable, including dummy variables for each customer mechanically results in a negative coefficient on the lagged dependent variable (see Nickell, 1981). Anderson and Hsiao (1981) were the first to show that the problem with the within estimator can be solved by differencing in order to eliminate the unobserved effect. Instrumental variables can then be used on the differenced variables in order to estimate unbiased coefficient values.[12]

Estimating dynamic models with an unobserved effect has proven to be more challenging for nonlinear models such as the case of binary response. No transformation has been found that is able to consistently eliminate the fixed effect in the same way as the Anderson-Hsiao procedure for the linear case. Thus, two fundamental challenges that arise include assumptions regarding the distribution of the unobserved individual effects and assumptions regarding the initial conditions of the dynamic process (Heckman, 1981). In this paper, we use a semiparametric method for estimating dynamic, binary-response models originally proposed by Cox (1958) and Chamberlain (1985) and more recently studied by Honore and Kyriazidou (2000). Unlike random-effects estimators, this fixed-effects method imposes less structure on the estimation. Most notably, it requires no assumptions to be made on the initial conditions of the process or on the distribution of the unobserved effects. This method has been employed by researchers in different settings to test for state dependence.[13]

If four or more observation periods are available for each individual, it is possible to identify first-order state dependence while controlling for unobserved effects.

[12] Appendix Table A provides results from using the linear models on our data.
[13] See for example Chay, Hoynes, and Hyslop (2001) for an analysis of state dependence in monthly welfare participation sequences.

Specifically, Cox (1958) showed that if the random disturbances are logistically distributed i.i.d., there exists a set of sufficient statistics $B \equiv \{y_{i1}, y_{iT}, s\}$, where $s = \sum_{t=1}^{T} y_{it}$, that can absorb both the individual effects and the initial conditions. Thus for the logit model,

(8) $$P(\mathrm{y}_{\mathrm{it}} \mid \alpha_i, y_{i1}, \ldots, y_{it-1}) = \frac{\exp(\gamma\, \mathrm{y}_{\mathrm{it\text{-}1}} + \alpha_i)}{1 + \exp(\gamma\, \mathrm{y}_{\mathrm{it\text{-}1}} + \alpha_i)},$$

the following conditional probability can be specified

(9) $$P(\mathrm{y}_{i1}, \ldots, y_{iT} \mid B) = \frac{\exp(\gamma \sum_{\mathrm{t=2}}^{\mathrm{T}} y_{it} y_{it-1})}{\sum_{d \in B} \exp(\gamma \sum_{\mathrm{t=2}}^{\mathrm{T}} d_t d_{t-1})}.$$

Note that the conditional probability does not depend on the parameter, α_i. Furthermore, conditioning on y_{i1} and y_{iT} solves the problems associated with the initial conditions.[14]

The intuition for this result is simple. Within a sufficiency class and in the absence of first-order state dependence, we would expect all sequences of events to occur with equal probability. The parameter γ will be estimated to be different than zero when certain sequences occur more frequently in the data than others of the same sufficiency class. For example, when T = 4, γ is identified by examining the following pairs of sequences: 1100 vs. 1010 and 0011 vs. 0101 where 1 represents a late-fee-paid visit and 0 represents a visit with no late fee paid. Notice that the unobserved effects are controlled for because the same number of 1's and 0's occur in each sequence. Furthermore, initial conditions are controlled for by comparing sequences with the same

[14] Incidentally, controlling for the initial and final conditions also controls for any problems with selective attrition in the sample.

starting and ending values. The only difference between these sequences is the "path" that is taken between the initial and final points. First-order state dependence suggests that 1010 and 0101 will occur more frequently in the data than 1100 and 0011 respectively. An estimate of γ can be obtained by maximizing the sample log-likelihood analog of Equation (9). Similar intuition holds when comparing sequences with more than four observations.

Chamberlain (1985) derives an estimator for second-order state dependence when at least 6 observation periods are available for each individual. If the random disturbances are logistically distributed i.i.d., The set of sufficient statistics is $B \equiv \{y_{i1}, y_{i2}, y_{iT-1}, y_{iT}, s, s_{11}\}$, where $s_{11} = \sum_{t=1}^{T} y_{it} y_{it-1}$. Thus for the logit model,

$$(10) \qquad P(\mathrm{y_{it}} \mid \alpha_i, y_{i1}, \ldots, y_{it-1}) = \frac{\exp(\gamma_1 \, \mathrm{y_{it\text{-}1}} + \gamma_2 y_{it-2} + \alpha_i)}{1 + \exp(\gamma \, \mathrm{y_{it\text{-}1}} + \gamma_2 y_{it-2} + \alpha_i)},$$

the following conditional probability can be specified

$$(11) \qquad P(\mathrm{y}_{i1}, \ldots, y_{iT} \mid B) = \frac{\exp(\gamma_2 \sum_{\mathrm{t=3}}^{\mathrm{T}} y_{it} y_{it-2})}{\sum_{d \in B} \exp(\gamma_2 \sum_{\mathrm{t=3}}^{\mathrm{T}} d_t d_{t-2})}.$$

It is noteworthy that this conditional probability does not depend on either α_i or γ_1. The intuition for this conditional probability is similar to that described above for testing first-order state dependence. When T = 6, the following pairs of sequences give conditional probabilities that contribute to the estimation of γ_2: 101000 vs. 100100, 000101 vs. 001001, 010111 vs. 011011, and 111010 vs. 110110. All of these pairs fall within the same sufficiency class and thus control for initial conditions and the unobserved individual-specific effects in the model. Second-order negative state dependence predicts that the second sequence in each of these pairs will occur more frequently in the data than the first.

For our analysis, we generate sequences of six observations so that both first-order and second-order state dependence can be estimated. This data set is created by extracting the first six observations for each movie-rental customer and then continuing to extract the subsequent six observations for each customer provided that six additional observations exist. After obtaining these sequences, the data set is further restricted to include only the 44 sequences of six observations which are useful for the testing of state dependence. This procedure leaves us with 7650 usable sequences of six observations. These sequences represent movie-rental behavior for 2,735 distinct customers.

Table 2 presents counts for each of the 44 different sequences used to test for first-order state dependence. The sequences are spaced such that each group represents a sufficiency class. Under the null hypothesis of no state dependence, the number of times that each sequence appears in the data should be statistically equivalent to all other sequences in the same sufficiency class. A quick comparison of the counts for sequences within a sufficiency class suggests that negative state dependence is present in this data.

V. Results

Column (1) of Table 3 provides the estimate of first-order state dependence obtained by maximizing the sample log likelihood analog of Equation (9) with respect to γ using the 7650 usable sequences. An estimate of $\gamma = -.1067$ is obtained from this procedure. This estimate is statistically significant at the 1% level and provides evidence in favor of punishment-induced deterrence. Since there are no other explanatory variables in the model, this coefficient has an easy to interpret partial effect (OLS-like interpretation). The coefficient suggests that an individual is 2.66% less likely to pay a late fee during visit t if a late fee was paid during visit t-1. Given that late fees are paid 14% of the time, receiving a late fee causes a 19% reduction in the probability of an individual receiving a late fee in the next period. These results are very similar to those presented in Appendix Table A which provides the results from a linear regression analysis.

Column (2) presents the estimate of second-order state dependence generated by maximizing the sample log likelihood analog of Equation (11) with respect to γ_2 using the 1648 sequences described earlier. An estimate of $\gamma_2 = -.0510$ is obtained using this procedure. This estimate suggests that paying a late fee during visit t-2 decreases the probability of paying a late fee during visit t by 1.27%. However, given the small sample of sequences that is usable to test for second-order state dependence, this effect while suggestive is not significantly different from zero at conventional confidence levels.

An important question regarding punishment-induced deterrence is if large penalties have a larger deterrent on future behavior than small penalties. In Columns (3) and (4) of Table 3 we attempt to address this issue. In Column (3) we reduce the sample to sequences for which the first late fee in the sequence was \$1 - \$3 (usually caused by returning one movie past the deadline by one day). Column (4), on the other hand, reduces the sample to sequences for which the first late fee in the sequence was greater than \$3 (usually caused by returning one movie past the deadline by more than one day or returning multiple movies late). We further restrict the samples in Columns (3) and (4) to be sequences of types (1) – (14) in Table 2. These are sequences of six observations for which there were two late fees. The reason for this restriction is that sequences in other sufficiency classes that test for first-order state dependence (e.g. 111000 vs. 110100) may not depend on the late fee amount in the first period. The sequences with exactly two late fees, however, all rely on the amount of the first late fee in the sequence and its effect on deterring subsequent late fee behavior.

The average and median late fee paid in our data conditional on the paid late fee being greater than \$3 is \$8.24 and \$6, respectively. Thus, the punishment meted out to individuals whose data are used in Column (4) is oftentimes several times larger than that given to the individuals whose data are used in Column (3). The results indicate that the punishment-induced deterrent effect of late fees greater than \$3 is almost twice as large ($\gamma = -.1313$) as the punishment-induced deterrent effect of late fees between \$1 and \$3 ($\gamma = -.0775$). This is consistent with the idea that larger late fees are more salient, implying greater punishment-induced deterrent effects.

We are interested in analyzing the effect of experience on punishment-induced deterrent effects. Given that many psychologically biases have been shown to disappear with experience, it is possible that the results from this analysis are being driven by movie-store customers who are still learning about their own late-fee proclivities or the late-fee policy of this particular video store. While it is likely that the majority of the movie-rental customers in our data set have been renting videos with this company for some time, we cannot rule out the possibility that some of the observations are people who started renting at the same time as our dataset began. In addition, we cannot rule out that these customers are learning information about themselves by receiving a late fee for the first time.

To address these experience issues, we first reduce the sample to sequences of six observations for people who, using our data, we can claim are experienced renters. In Columns (1) – (3) of Table 4, we restrict the sample to sequences for which the customer had previously rented at least 10, 20, and 40 times, respectively. Using these restricted samples, we estimate the level of first-order negative state dependence in the data. The results indicate that punishment-induced deterrence is just as strong (if not stronger) when restricting the data to individuals who could be considered experienced renters. Columns (4) – (6) present similar estimates when reducing the sample to sequences where the renter had previously paid at least 2, 4, or 10 late fees. Once again we find strong evidence in favor of punishment-induced deterrence.

While Columns (1) – (6) of Table 4 indicate that even people who have significant experience with renting movies exhibit punishment-induced deterrence, these estimates only provide insight into how experience might mitigate punishment-specific deterrence behavior under the assumption that γ is homogeneous across the population. In fact, all of the estimates in Columns (1) – (6) being larger than the full sample estimate suggests that individuals that rent a lot of movies are more susceptible to this bias than those that rent only a few movies. If this is the case, when we restrict the sample to experienced renters, we are also restricting the sample to people with a higher γ and therefore not testing for the mitigating effects of experience. In order to correctly identify the effects of experience given customer heterogeneity, we split the sequences of each individual in half.[15] Columns (7) and (8) present the results when the data is restricted to the first half and second half of sequences for each individual, respectively.

[15] We drop the last sequence of 6 observations for each individual who has an odd number of sequences in the data.

The resulting estimates of first-order state dependence provide γ estimates for the first half and second half of sequences to be -.1493 and -.1118. While the estimate for the second half is slightly lower than the first half (evidence in favor of experience reducing the amount of punishment-induced deterrence), both estimates are still significant but the difference between the estimates is not significant.

Paying a late fee might have a deterrent-type effect on customers in other dimensions aside from reducing future late fees. We test to see if individuals who pay a late fee decide not to visit the video store as often and/or decide to rent fewer movies in subsequent visits. In Table 5 we present the results from the following specification

$$\text{(12)} \qquad \mathrm{Y_{it}} = \alpha_i + \gamma_1 \text{ Paid Fee}_{\text{it-1}} + \gamma_2 \text{ Paid Fee}_{\text{it-2}} + \gamma_3 \text{ Paid Fee}_{\text{it-3}} + \mu_{it}$$

where Y_{it} represents either the number of days between visit t and t-1 or the number of movies rented during visit t. Since a lagged dependent variable does not enter into the model anymore, we are able to use fixed effects to control for individual heterogeneity. Since the dependent variables (days between rentals and movies rented) are both counts, we present fixed effects results from both OLS and Poisson models.

Columns (1) and (2) of Table 5 suggest that after controlling for unobserved individual heterogeneity, paying a late fee is associated with an individual waiting 0.73 additional days before returning to the video store to rent another movie. This result appears to decay quickly over time. Paying a late fee two periods ago continues to be associated with a statistically significant longer waiting time before returning to the video store (0.48 days). However, paying a late fee three visits ago does not have a statistically significant effect on the number of days between rentals. Columns (3) and (4) of Table 5 test whether paying a late fee reduces the number of videos that the customer will rent during the subsequent visit. While the point estimates are all negative (fewer videos rented after paying a late fee), we find no significant evidence that the number of videos rented is reduced.

VI. Conclusion

While criminologists have long considered the possibility that being punished for a crime can serve as a mechanism to reduce repeat offenses, disentangling punishment-induced deterrence from classical deterrence in actual crime settings has proven to be very difficult. This analysis abstracts from actual criminal decisions by using data from the video-rental market in order to obtain an accurate measure of punishment-induced deterrence. We find evidence that video-return decisions are consistent with the hypothesis of punishment-induced deterrence. Our results further suggest that the deterrence caused by having to pay a fine in a previous period is increasing with respect to both the size and temporal proximity of the fine.

The key results found in this paper are statistically significant; however, it is important to discuss whether they are economically large. Our results indicate that while the immediate effect of punishment-induced deterrence is large (19% reduction in the subsequent visit), it decays quickly over time. It is important to remember, however, that this effect is being caused by very small fines (60% of the late fees are $3 or less). Another way to think about the size of the punishment-induced deterrent effect is the impact of this behavior on firm revenue. Using just the coefficient on first-order state dependence, we estimate that punishment-induced deterrence reduced late-fee revenue received during the two years for which we have data available by approximately $5,000 for this video store.

While the results from our analysis have direct implications regarding consumer rationality and marketing, caution should of course be taken when extending the analysis to crime. Clearly, serious criminal behavior is much more complex than returning movies late; it may be reasonable, however, to expect our results to be generalizable to less serious crimes that give fines as punishments (e.g. traffic and parking tickets). This analysis sacrifices some external validity in order to accurately identify punishment-induced deterrence. Having found evidence that punishment-induced deterrent effects can be sizable, we hope that future research will be able to more fully understand the impact that experiencing punishments can have on criminal behavior.

References

Anderson, T.W., and Hsiao, C., "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, 76,1981, 598-606.

Barber, Brad M. and Terrance Odean. "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors", mimeo, 2002.

Becker, Gary S., "Crime and Punishment: An Economic Approach," The Journal of Political Economy, Mar.-Apr. 1968, Vol 76 No. 2, 169-217.

Bentham, Jeremy, The Works of Jeremy Bentham, Vol. 1, edited by Bowning. New York: Russell and Russell, [1843] 1962.

Case, A. and L. Katz. "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths", NBER Working Paper No. 3705, May 1991.

Chay, Kenneth Y., Hoynes, Hilary, and Dean Hyslop. "A Non-experimental Analysis of True State Dependence in Monthly Welfare Participation Sequences," mimeo, 2001.

Chen, Keith M. and Jesse M. Shapiro, "Does Prison Harden Inmates? A Discontinuity-based Approach", Mimeo, 2005.

Clarke, Ronald V., "Approved School Boy Absconders and Corporal Punishment," British Journal of Criminology, 1966, 6, 364-375.

Chamberlain, G. "Heterogeneity, Omitted Variable Bias, and Duration Dependence," in Longitudinal Analysis of Labor Market Data, edited by J.J. Heckman and B. Singer. Cambridge: Cambridge University Press, 1985.

Cohen, Lawrence E. and Rodney Stark, "Discriminatory Labeling and the Five-finger Discount," Journal of Research in Crime and Delinquency, 1974, 11, 25-39.

Cox, D.R. "The Regression Analysis of Binary Sequences," *Journal of the Royal Statistical Society,* Series B, 20, 1958, 215-232.

DellaVigna, Stefano and Joshua Pollet, "Attention, Demographics, and the Stock Market", mimeo, 2006.

Di Tella, Rafael and Ernesto Schargrodsky, "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack," American Economic Review, March 2003, 94(1), 115-133.

Farrington, David P., "The Effects of Labeling," British Journal of Criminology, 1977, 17, 112-25.

Fiske, Susan T. and Shelley E. Taylor, "Social Cognition", New York: McGraw-Hill, Inc., 1991.

Glaeser E., B. Sacerdote and J. Scheinkman. "Crime and Social Interactions", Quarterly Journal of Economics, CXI, 1996, 507-548.

Glueck, Sheldon, "Principles of a Rational penal Code," Harvard Law Review, 1928, 41, 453-82.

Gold, Martin and Jay Williams, "National Study of the Aftermath of Apprehension," 1969, Prospectus 3, 3-11.

Heckman, J.J., "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process," Chapter 4 in C. Manski and D. McFaden (eds.), Structural Analysis of Discrete Data, MIT Press, Cambridge, MA, 1981.

Hong, Harrison, Walter Torous, and Ross Valkanov. "Do Industries Lead the Stock Market?" mimeo, 2003.

Honore, Bo E. and Ekaterini Kyriazidou. "Panel Data Discrete Choice Models with Lagged Dependent Variables", *Econometrica,* Vol. 68, No. 4. (Jul., 2000), pp. 839-874.

Huberman, Gur, and Tomer Regev. "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar", *Journal of Finance,* 2001, Vol. 56, 387-396.

Jolls, Christine, Sunstein, Cass R., and Richard H. Thaler, "A Behavioral Approach to Law and Economics", *Stanford Law Review,* 50, 1998, 1471-1550.

Katz, L., A. Kling, and J. Liebman. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment", Quarterly Journal of Economics, 2001, CXVI, 607-654.

Katz, Lawrence F., Steven D. Levitt, and Ellen Shustorovich, "Prison Conditions, Capital Punishment and Deterrence," American Law and Economics Review, Fall 2003, 5(2), 318-343.

Kessler, Daniel and Steven D. Levitt, "Using Sentence Enhancements to Distinguish Between Deterrence and Incapacitation," Journal of Law and Economics, April 1999, 42 (1, Part 2), 343-363.

Klemke, Lloyd W., "Does Apprehension for Shoplifting Amplify or Terminate Shoplifting Activity," Law and Society Review, 1978, 12, 391-403.

Lee, David S. and Justin McCrary, "Crime, Punishment, and Myopia", NBER Working Paper #11491, 2005.

Levitt, Steven D., "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation," Quarterly Journal of Economics, May 1996, 111(2), 319-351.

Levitt, Steven D., "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," American Economic Review, June 1997, 87(3), 270-290.

List, John A., "Does Market Experience Eliminate Market Anomalies?," *Quarterly Journal of Economics* (2003), 118(1), pp. 41-71.

List, John A., "Neoclassical Theory Versus Prospect Theory: Evidence From the Marketplace," *Econometrica*, 2004, v72(2, Mar), 615-625.

McCord, Joan, "A Longitudinal Appraisal of Criminal Sanctions," Paper presented to the IXth International Congress on Criminology, Vienna, Austria, September 1983.
Schwartz, Richard and Jerome Skolnick, "Two Studies of Legal Stigma," Social Problems, 1962, 10, 133-42.

Nickell, S. (1981). "Biases in Dynamic Models with Fixed Effects." *Econometrica*, 49, 1981, 1417-1426.

Pintoff, Randi, "Juvenile Jails: A Path to the Straight and Narrow or Hardened Criminatlity?," Mimeo, 2005.

Polinsky, A. Mitchell and Steven Shavell, "The Optimal Tradeoff Between the Probability and Magnitude of Fines", American Economic Review, Dec. 1979, Vol. 69 No. 5, 880-891.

Polinsky, A. Mitchell and Steven Shavell, "A Note on Optimal Fines When Wealth Varies Among Individuals", American Economic Review, June 1991, Vol 81, No. 3, 618-621.

Sah, R. "Social Osmosis and Crime", Journal of Political Economy, 1991, 99, 1272-1295.

Sherman, L.W. and R. Berk. "The Specific Deterrent Effects of Arrest for Domestic Assault." *American Sociological Review,* 49 (1984): 261-272.

Shoham, S. Giora, "Punishment and Traffic Offenses," Traffic Quarterly, 1974, 28, 61-73.

Smith, Douglas A. and Patrick R. Gartin, "Specifying Specific Deterrence: The Influence of Arrest on Future Criminal Activity," *American Sociological Review*, 1989, 54:94-106.

Table 1. Summary Statistics - By Individual

	Mean	Standard Deviation	Median	Min	Max
Visits (2-year period)	21.4	29.6	9	1	320
Avg Movies Rented (per visit)	2.3	1.1	2	1	12
Fraction of Time Movies are Returned Late	0.14	0.20	0.04	0	1
Late Fees Paid ($, per visit, conditional on paying a late fee)	4.24	3.34	3.3	1	44
Late Fees Paid ($, 2-year period)	16.5	45.1	2	0	1335
Total Number of Customers	10563	10563	10563	10563	10563

Notes: Summary statistics represent data from all video-store transactions made between Jan. 1, 2003 – Dec. 31, 2004. A visit represents all transactions that take place on a given day by a customer account number.

Table 2. Counts of Different Sequence Types Used For Testing First-order State Dependence

(1)	110000	266	(27)	011100	114
(2)	101000	307	(28)	001110	117
(3)	100100	317	(29)	010110	146
(4)	100010	288	(30)	011010	149
(5)	000011	287	(31)	111100	59
(6)	010001	322	(32)	111010	74
(7)	000101	339	(33)	110110	82
(8)	001001	345	(34)	101110	85
(9)	011000	300	(35)	001111	87
(10)	001100	330	(36)	011101	75
(11)	000110	341	(37)	010111	83
(12)	001010	328	(38)	011011	101
(13)	010010	346			
(14)	010100	347	(39)	100111	71
			(40)	110011	80
(15)	111000	103	(41)	111001	82
(16)	110100	120	(42)	110101	70
(17)	110010	123	(43)	101101	77
(18)	100110	125	(44)	101011	100
(19)	101100	128			
(20)	101010	137			
(21)	000111	123			
(22)	001011	112			
(23)	010011	135			
(24)	011001	137			
(25)	001101	138	**Total No.**		
(26)	010101	154	**of Sequences:**		7650

Notes: Each sequence type represents six consecutive visits by the same individual. 1's indicate that a late fee was paid during that visit and 0's indicate no late fee paid. Types (1) – (44) illustrate all sequences of six visits that are usable to test for first-order state dependence. Sequence types are separated into groups ((1)-(4), (5)-(8), etc.) which represent a given sufficiency class. The third and sixth columns provide counts for the number of times the sequence occurs in our data.

Table 3. Fixed-Effects Estimates of State Dependence - Based on Semiparametric Conditional Logit Models

	Dependent Variable: Paid Fee in Period (t)			
	(1)	(2)	(3)	(4)
Paid Fee (t-1)	-0.1067		-0.0775	-0.1313
	(.0237)***		(.0416)*	(.0499)**
Paid Fee(t-2)		-0.0510		
		(.0464)		
First of Two Paid Fees $1-$3			X	
First of Two Paid Fees > $3				X
Log Likelihood	-18661	-1142	-6638	-3633
Total No. Observations	45900	9888	16614	9216
Total No. Chains of Six	7650	1648	2769	1536

Notes: Columns (1) – (4) provide maximum likelihood estimates of state dependence using the conditional log-likelihood functions given in Equations (9) and (11) – Equation (9) represents first-order state dependence and Equation (11) represents second-order state dependence. Standard errors are computed using a bootstrap routine with 1000 repetitions of full samples with replacement. Column (3) uses the subset of sequences which have exactly two late fees and where the first late fee paid is between $1 and $3. Column (4) uses the subset of sequences which have exactly two late fees and where the first late fee paid is greater than $3.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4. Estimating the Effects of Experience on First-Order State Dependence

	Dependent Variable: Paid Fee in Period (t)							
	Number of Previous Visits			Number of Previous Late Fees			First	Second
	>10	>20	>40	>2	>5	>10	Half	Half
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Paid Fee (t-1)	-0.1540	-0.1238	-0.2227	-0.1127	-0.1803	-0.1674	-0.1493	-0.1118
	(.0281)***	(.0327)***	(.0445)***	(.0284)***	(.0333)***	(.0411)***	(.0398)***	(.0386)***
Log Likelihood	-13451	-9859	-5456	-13620	-9736	-6010	-7131	-7157
Total No. Observations	33042	24300	13446	33690	24078	14784	17580	17580
Total No. Chains of Six	5507	4050	2241	5615	4013	2464	2930	2930

Notes: Columns (1) – (8) provide maximum likelihood estimates of state dependence using the conditional log-likelihood functions given in Equation (9) in the text. Standard errors are computed using a bootstrap routine with 1000 repetitions of full samples with replacement. Columns (1) – (3) restrict the sample by not creating sequences of six observations for each individual until the first 10, 20, and 40 visits to the video store have been deleted, respectively. Columns (4) – (6) restrict the sample by not creating sequences of six observations until the individual has paid 2, 5, and 10 late fees, respectively. Column (7) restricts the sample by only including the first half of sequences for any individual. Column (8) restricts the sample by only including the second half of sequences for any individual. In the event of an odd number of sequences for a given individual, the last sequence is deleted.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5. The Effect of Receiving a Late Fee on Time Between Rental Periods and Movies Rented Per Visit - OLS and Poisson Models

	Dependent Variable: Number of days between movie rental (t) and movie rental (t-1)		Dependent Variable: Number of movies rented during visit t	
	OLS	**Poisson**	**OLS**	**Poisson**
Late Fee (t-1)	0.732	0.051	-0.015	-0.006
	(.153)***	(.010)***	(.010)	(.003)*
Late Fee (t-2)	0.477	0.034	-0.009	-0.004
	(.150)***	(.010)***	(.010)	(.004)
Late Fee (t-3)	0.247	0.019	-0.017	-0.007
	(.152)	(.012)	(.010)	(.004)*
Individual F.E.	X	X	X	X
Observations	198,174	198,174	198,174	198,174

Notes: In Columns (1) and (2), the dependent variable is a count of the number of days between the current movie-rental visit (visit t) and the last time that the customer rented a movie (visit t-1). In Columns (3) and (4), the dependent variable is a count of the total number of movies that the customer rented in the current movie-rental visit (visit t). Columns (1) and (3) use ordinary least squares with customer fixed effects. Robust standard errors for these columns are presented in parentheses. Columns (2) and (4) run a Poisson conditional fixed effects model. Bootstrapped standard errors for these columns are presented in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix A. Estimates of State Dependence Based on the Linear Probability Model

	Linear Probability Model				IV (LPM)
	(1)	(2)	(3)	(4)	(5)
Late Fee Paid (t-1)	0.154	-0.023	-0.023	-0.027	-0.018
	(.003)***	(.003)***	(.004)***	(.005)***	(.004)***
Fraction Late (t-1, t-10)		0.693			
		(.007)***			
Fraction Late (t-1, t-25)			0.858		
			(.009)***		
Fraction Late (t-1, t-50)				0.926	
				(.014)***	
Adj. R-Squared	0.024	0.108	0.129	0.129	0.018
Observations	215,216	154,337	96,037	46,253	206,263

Notes: The dependent variable in Columns (1) – (4) is an indicator that equals one if the customer paid a late fee during that visit. Robust standard errors are presented in parentheses. Fraction Late (t-1, t-X) is a variable that equals the fraction of time that the customer paid a late fee in the previous X visits. Column (5) uses the Anderson-Hsiao method with the dependent variable being the difference between the late-fee-paid indicator in period t and the late-fee-paid indicator in period t-1. The Late Fee Paid (t-1) difference is instrumented with Late Fee Paid (t-2).
* significant at 10%; ** significant at 5%; *** significant at 1%