Modeling Sample Selection for Durations with Time-Varying Covariates, With an Application to the Duration of Exchange Rate Regimes

Frederick J. Boehmke
University of Iowa
Department of Political Science

Christopher M. Meissner
University of California, Davis and NBER
Department of Economics

September 30, 2009

Contact: frederick-boehmke@uiowa.edu. Boehmke is Associate Professor of Political Science at the University of Iowa. Meissner is Associate Professor of economics at the University of California, Davis and a Research Associate at the NBER. We thank Neal Beck, Michael Colaresi, Brad Jones, and Michael Klein for thoughtful comments and suggestions.

Abstract

We extend existing estimators for duration data that suffer from non-random sample selection to allow for time-varying covariates. Rather than a continuous-time duration model, we propose a discrete-time alternative that models the effects of sample selection at the time of selection across all subsequent years of the resulting spell. Properties of the estimator are compared to those of a naive discrete duration model through Monte Carlo analysis and indicate that our estimator outperforms the naive model when selection is non-trivial. We then apply this estimator to the question of the duration of monetary regimes and find evidence that ignoring selection into pegs leads to faulty inferences.

1 Introduction

The consequences of non-random sample selection have been known among political scientists for quite some time. A growing body of literature exists that documents the consequences of ignoring sample selection and that demonstrates its effects on our understanding of real-world political phenomena, including voter turnout, interest group lobbying, public opinion, and the outcome of international crises. As the methodological tools and interests of political scientists have developed, we have extended our studies into situations where existing techniques for correcting selection bias do not fit the question in hand. Specifically, the last decade has seen a dramatic rise in the use of duration models to explain the time until political events occur, including regime transitions, the confirmation of political nominees, position-taking by elected representatives and the duration of cabinets in parliamentary democracies. Unfortunately, until recently there has been no way to deal with issues of sample selection in these and other duration analyses, despite many theoretical advances that indicate that selection should be a concern.

In response to this gap, researchers have proposed a couple of approaches for dealing with sample selection issues in the duration context. Prieger (2002) uses copulae (see, e.g., Smith 2003) to bind together two marginal distributions while Boehmke, Morey and Shannon (2006) use bivariate distributions to accomplish the same task. The latter demonstrates through Monte Carlo analysis that ignoring sample selection issues can result in biased parameter estimates when estimating naive duration models — including the exponential, Weibull, and Cox — on data that suffer from selectivity.

While political scientists are already applying these estimators (e.g., Beardsley and Asal 2009; Long, Nordstrom and Baek 2007), one significant shortcoming is that they do not allow for time-varying covariates, despite the fact that many, perhaps most, applications of duration models involve explanatory factors that change over the course of a single spell. Unfortunately, extending existing models to permit time varying covariates are complicated by the move from a single stochastic component for the entire duration to a series of them — one for each interval of a spell (e.g., each year or day). Even without the sample selection component, any (parametric) duration model that allows for time-varying covariates (hereafter, TVCs) models the probability of failure within each interval of a spell, rather than the continuous hazard at each point in time.

Given this difference between estimators for time-invariant covariates and those for TVCs, developing an estimator that accounts for sample selection in durations with TVCs requires a different approach than that used for continuous-time durations without TVCs. In this paper we develop such an estimator. Our approach reflects the discrete nature of failure within an interval by joining probit models for selection and duration. While moving to a discrete duration model means that we do not model duration dependence directly through the parametric shape of the error distribution (e.g., Weibull, log-normal), duration dependence is still easily modeled through the inclusion of appropriate variables measuring time since a spell begins. We allow the error in the selection equation from the year the spell begins to be correlated with the errors in each interval for the subsequent spell; we also assume that despite this correlation, errors within a spell are uncorrelated with each other. In order to permit greater flexibility, we allow the correlation to decay over time so that the effect of selection may decrease over the course of a spell as an observation moves farther away from the circumstances at the time of selection. This assumption both extends the applicability of our estimator and also likely fits with empirical reality: the conditions that help engender selection will generally have a large effect early on in the ensuing duration process, but their effect will often decrease over time as circumstances change and new forces swamp the initial conditions.

In order to evaluate the usefulness of our proposed estimator, we perform a series of Monte Carlo simulations comparing its results to those obtained from naive discrete-time duration estimators. We vary both the initial correlation as well as its rate of decay. Our results provide strong support for our new estimator, which outperforms the naive estimator in terms of bias and root mean squared error whenever the correlation is non-trivial.

We then apply our estimator to the study of the duration of monetary regimes. A recent view holds that the decline in the average duration of pegged exchange rates in the early 1990s was likely caused by the increasing global integration of capital markets (Obstfeld and Rogoff 1995). We argue that the kind of countries that adopt pegs differ from the general population of countries. In fact, our correction for sample selection finds evidence consistent with the idea that unobserved factors that lead a country to choose a peg also make that country more likely to drop out of that peg. Such evidence is consistent with the idea that countries use only short-term criteria to judge the benefits and costs of pegs or systematically underestimate their ability to maintain a peg. Further, correcting for non-random sample selection also changes our inferences about key explanatory factors. In particular, we find a greater role for political factors such as political stability, and a reduced role for economic factors such as reserve accumulation. These findings are consistent with the idea that reserves may merely be a signal of the resolve of a country to defend a peg rather than being the means to the end of defense of the peg.

2 Existing Estimators for Duration Data with Selection

Non-random sample selection is a problem for standard estimators because unobserved factors that influence the duration (or, more generally, the quantity) of interest also influence whether or not that observation makes it into the sample at all. When this relationship exists, the value of the dependent variable of interest is related to the selection process, since it depends on both the systematic and the unobserved stochastic components. Observed values of the dependent variables are therefore not representative, even for the observed sample. Because the dependent variable is unrepresentative, parameter estimates are biased, even after controlling for individual characteristics through independent variables. The severity of the problem depends on the correlation between explanatory factors in the selection and outcome equations. When they are correlated, this generally induces correlation between the error term in the equation of interest and the independent variables. Thus an additional assumption is violated and the bias is generally exacerbated. In more complicated models, bias in one parameter can lead to bias in other parameters.

A common solution for selection bias involves modeling the selection process and estimating its parameters while simultaneously estimating the parameters of the equation of interest (e.g., Heckman 1976 and 1979; Dubin and Rivers 1990). In these cases, one conditions on the selection process when estimating the quantity of interest. With a properly specified selection equation, these estimators generally produce consistent parameter estimates.

Boehmke, Morey and Shannon (2006) build on these results by developing an estimator for duration data with possible non-random sample selection. They use a bivariate exponential distribution to link the discrete outcome of the selection equation with the continuous duration outcome of interest. This estimator is then extended to allow for Weibull duration dependence. The derivation parallels that used to correct for selection bias with a continuous (Heckman 1976, 1979) or a discrete (Maddala 1983; Dubin and Rivers 1990) dependent variable. By jointly modeling the selection and duration processes, consistent estimates are obtained. An alternate approach is taken by Prieger (2002), which uses copulae to combine a probit selection equation with a Weibull duration equation.

In these estimators, the duration component corresponds to a continuous-time duration model with time-invariant covariates.[1] Yet many applications of duration models involve TVCs. For example, the duration of cabinets may depend on economic performance, which changes from month-to-month or year-to-year; the duration of militarized interstate disputes may depend on the losses taken by each side or by the actions of third parties that try to intervene. Standard continuous-time duration models can be easily modified to allow for TVCs by partitioning each spell into intervals during which included covariates do not change. These intervals may be days, months, or years depending on the frequency with which observed values change. The likelihood of each spell is then calculated with the product of the probability of surviving each interval given survival until that interval, until the last period, which contributes either a discrete probability of failure if failure is observed or a discrete probability of survival if the observations is right-censored.[2]

In essence, then, estimating a continuous-time duration model with TVCs is quite the same as estimating a discrete-time duration model. The two main differences arise from the distributional assumptions regarding the error terms — parametric continuous time models often assume a Weibull distribution whereas discrete ones assume a logistic or normal distribution — and the treatment of duration dependence, which is accomplished implicitly through the distribution in continuous models and explicitly through the possible inclusion of covariates relating to time in the discrete-time models.

The move to TVCs therefore changes the structure of the estimation in such a way as to make it difficult to apply existing solutions for sample selection with duration data. Partitioning each spell into different components and then calculating the probability of failure in each interval changes

---

[1] See Box-Steffensmeier and Jones (2004) for more information on duration models in general.

[2] If the precise moment of failure is observed, then the last interval contributes the density of the time of failure given survival until that interval.

the data generating process from one in which each spell has a single stochastic component to one in which each interval of each spell has its own stochastic term. This makes it impossible to directly apply existing solutions for non-random sample selection in duration models. Further, it makes it difficult to extend existing solutions since one must correlate the unobserved terms at the time of selection with potentially more than one unobserved term for each interval in the duration of the subsequent spell. This leads us to propose an alternate, though related, form of the estimator that allows for both sample selection and time-varying covariates.

3 Modeling Sample Selection in Discrete-Time Durations

Given the complications just outlined and the discrete nature of the duration process, we move from previous estimators' use of continuous-time duration models to discrete-time duration models. As with previous estimators, we continue to model the selection process as a discrete outcome, but also model the duration process as a discrete outcome. Once an observation has selected into the duration process, then, we assume that we observe a discrete indicator for failure for each interval of the corresponding spell. This estimator therefore has a lot in common with discrete versions of the Heckman model (Dubin and Rivers 1990; Maddala 1983), but rather than observing a single outcome (e.g., whether an individual registers to vote and whether registered individuals turn out on election day), we observe a series of zeros for each interval the individual survives followed by a single one corresponding to failure in the last interval of the spell.

In the standard extension of the Heckman model for dichotomous outcomes, the two stochastic components are allowed to have non-zero correlation in order to capture possible non-random sample selection. This involves a straightforward application of, for example, the bivariate normal distribution. If all of our durations were observed for only one interval, we too could apply this estimator (though we would not have TVCs if that were the case). But because our duration is measured as a vector of zeros followed by a one (or a terminal zero in the case of right-censoring), there are a series of stochastic terms — one for each period of the ongoing spell — that could possibly be correlated with unobserved components at the time of selection. This necessitates modifying the standard dichotomous selection estimator to account for the duration structure.

In the following section we propose such an estimator by allowing the stochastic term in the selection equation at the time of entry to be correlated with the stochastic terms for each interval of the resulting duration. Basically, this means that the effects of unobserved variables influencing selection persist over the entire duration of the ensuing spell, but that errors across periods in a spell are not correlated. This leads to an estimator that involves a combination of a sample selection model for a discrete outcome to account for the first period of a spell with the same model for stochastic truncation in subsequent periods.

4 Derivation of the Likelihood Function

In order to derive the full likelihood, we first describe the selection and duration equations separately and then discuss how we link the two to account for possible non-random sample selection. Both equations are represented with standard binary outcome models.

For the selection equation, let C_i indicate whether an individual, i, selects into the duration sample and let the probability that $C_i = 1$ depend on some vector of covariates W_i. Assume a standard threshold model as follows:

$$\begin{aligned} C_i^* &= W_i\gamma + \eta_i; && (1) \\ C_i &= \begin{cases} 1 & \text{if } C_i^* > 0, \\ 0 & \text{otherwise.} \end{cases} && (2) \end{aligned}$$

Ultimately, we will use the bivariate normal to link the unobserved components, which makes the selection equation a standard probit model. This set up is identical to that used in most selection estimators.

The duration equation is modeled in a similar way, but since we must model duration in each interval of a given spell, we add a subscript for time, t. The length of time represented by each increment of t is determined by the largest unit of time such that covariates are constant within that interval. In most political science applications, t will represent days, months or years. Failure is measured by the binary variable Y_{it}, which indicates whether individual i's duration ends at time t. Let T_i correspond to the final interval of i's spell. Thus Y_{iT_i} equals one and all other realizations of Y_{it} are zero.[3] We assume that the duration depends on a vector of covariates, X_{it}, at least some of

---

[3] Right-censoring is handled trivially by this model. Rather than ending with a one, right-censored observations merely end with a zero in the last interval when censoring occurs.

which vary over intervals of time. Duration is therefore modeled using a discrete outcome model, which is common for discrete event history analyses (see, e.g., Box-Steffensmeier and Jones 2004):

$$\begin{aligned} Y_{it}^* &= X_{it}\beta + \epsilon_{it}; && (3) \\ Y_{it} &= \begin{cases} 1 & \text{if } Y_{it}^* > 0, \\ 0 & \text{otherwise.} \end{cases} && (4) \end{aligned}$$

Sample selection occurs when the duration data are not observed for observations with $C_i = 0$. This means that all values of Y_{it} for an entire spell go unobserved.

To put these two pieces together into a single estimator, we partition the data into three groups. The first group involves all cases that do not select in, the second group represents the first interval of a duration spell for cases that select in, and the third group represents all additional intervals of the spell for observations that have selected in and survived the first interval. We partition the data in this way to facilitate construction of the likelihood. At the moment of selection, observations contribute two pieces of information: that they select in and whether they survive the first interval. Observations that survive the first interval are already known to have selected in and this information has already been incorporated into the likelihood, so they only contribute information about whether they survive additional intervals given that they have already selected in.

Put into probability statements, then, the three pieces are: $\Pr(C_i = 0|W_i)$, $\Pr(C_i = 1, Y_{i1} = y_{i1}|W_i, X_i)$, and $\Pr(Y_{it} = y_{it}|W_i, X_i, C_i = 1, 1 < t)$. The likelihood of the data can be written out as the product of these three components:

$$\begin{aligned} \Pr(\mathbf{Y}, \mathbf{C}) &= \prod_{C_i=0} \Pr(C_i = 0|W_i) \times \prod_{C_i=1} \Pr(C_i = 1, Y_{i1} = y_{i1}|W_i, X_{it}) && (5) \\ &\quad \times \prod_{C_i=1, t>1} \Pr(Y_{it} = y_{it}|W_i, X_{it}, C_i = 1). && (6) \end{aligned}$$

Taken separately, the first two terms constitute a selection estimator for discrete outcomes that includes data for observations that select in as well as those that do not. The third term represents a selection estimator for stochastic truncation, which includes only information about individuals that select in. Our likelihood is therefore a combination of two commonly used estimators that individually represent the different types of information that an observation contributes.

Once assumptions are made about the distribution of the error terms and the functional form of the estimator, these densities and probabilities can be explicitly calculated and a likelihood function can be specified. Here, we assume that the stochastic terms are generated according to a bivariate normal distribution with correlation $\rho_t = Corr(\eta_i, \epsilon_{it})$. This is a key assumption. First, as the subscript indicates, we allow the correlation to change over time. Second, we assume that the duration errors from different intervals of a spell are not correlated with each other. This second assumption means that selection bias is captured entirely through correlation of the error terms in each interval of time with the selection equation error term at the time of selection; no additional information is gained during the course of a duration. We make this restrictive assumption because to allow for correlation across the stochastic terms for each interval in a given spell would essentially involve a much more complicated time-series cross-sectional model with autocorrelation, which has proved difficult to estimate without restrictive assumptions.[4]

The first assumption of non-constant correlation is made for both substantive and statistical reasons. Constant correlation over time results in a limit on the maximal absolute correlation in order to maintain semi-positive definiteness of the covariance matrix for the selection equation error and the sequence of duration equations stochastic terms.[5] While all covariance matrices must meet this condition, the structure of the one used for this estimator means that satisfying it depends solely on the value of the correlation parameter. For example, this just means that when $T = 1$, giving a 2×2 covariance matrix, that the correlation can not be greater than one. With longer spells, however, the maximal correlation decreases and becomes quite low.

To allow more flexibility, then, we assume that the correlation decays exponentially over time. Substantively, this implies that unobserved factors in the selection process become less and less important over the course of spell. We do allow some flexibility in the decay process, however, by parameterizing it as follows: $\rho_t = \rho_0 \exp(-\delta(t-1))$, where ρ_0 describes the correlation between the errors from the selection equation and the first interval of the spell and $\delta \geq 0$ allows the rate of decay to vary. Figure 1 presents examples of the resulting correlation over time for different values of ρ_0 and δ. Note that when δ is large, the correlation goes to zero after only five periods, but when it is small, the correlation is still nontrivial after 15 periods. Of course, there is a tradeoff between the two parameters: large initial correlations will have to decay faster whereas small ones

---

[4] see Pang (2008) for a recent discussion and a promising Bayesian approach.

[5] When the correlation is constant over T periods the determinant of the covariance matrix is $1 - T\rho^2$, leading to an upper bound on the correlation of $|\rho| \leq \sqrt{1/T}$.

can persist for long periods of time.

[Figure 1 about here.]

While this parameterization extends the maximal spell length, it does so by assuming that the correlation decreases over time. We believe that this assumption has some intuitive appeal: while unobserved components that influence selection may a have a strong relationship with unobserved components early on in a spell, it seems reasonable to assume that that relationship will weaken over time as the conditions present at selection recede into the past and contemporaneous unobserved events take precedence. Even with the assumption, the maximal initial correlation still depends on the length of the observed spells, but now it also depends on the rate of decay. Specifically, with a maximal spell length of T, the restriction on the maximal initial correlation is

$$|\rho_0| \leq \sqrt{\frac{1-\exp(-2\delta)}{1-\exp(-2\delta T)}}. \tag{7}$$

Figure 2 displays this relationship for four different values of the decay parameter and different maximal spell lengths.

[Figure 2 about here.]

With the bivariate normal assumption, we can write out the corresponding likelihood by calculating each of the component probabilities, where $\Phi(z)$ represents the cumulative standard normal density and $\Phi(z_1, z_2, \rho_t)$ represents the cumulative bivariate standard normal density with correlation ρ_t at time t.

$$\begin{aligned}
\Pr(C_i = 0|W_i) &= \Pr(W_i\gamma + \eta_i \leq 0|W_i), && (8)\\
&= \Phi(-W_i\gamma); && (9)\\
\Pr(C_i = 1, Y_{i1} = 1|W_i, X_{i1}) &= \Pr(W_i\gamma + \eta_i > 0, X_{i1}\beta + \epsilon_{i1} > 0|W_i, X_{i1}), && (10)\\
&= \Pr(\eta_i > -W_i\gamma, \epsilon_{i1} > -X_{i1}\beta|W_i, X_{i1}), && (11)\\
&= \Pr(\eta_i \leq W_i\gamma, \epsilon_{i1} \leq X_{i1}\beta|W_i, X_{i1}), && (12)\\
&= \Phi(W_i\gamma, X_{i1}\beta, \rho_0); && (13)\\
\Pr(C_i = 1, Y_{i1} = 0|W_i, X_{i1}) &= \Phi(W_i\gamma, -X_{i1}\beta, -\rho_0); && (14)\\
\Pr(Y_{it} = 1|W_i, X_{it}, C_i = 1) &= \frac{\Pr(C_i = 1, Y_{i1} = 1|W_i, X_{it})}{\Pr(C_i = 1|W_i)}, && (15)\\
&= \frac{\Phi(W_i\gamma, X_{it}\beta, \rho_t)}{\Phi(W_i\gamma)}; && (16)
\end{aligned}$$

$$\Pr(Y_{it} = 0|W_i, X_{it}, C_i = 1) \;=\; \frac{\Phi(W_i\gamma, -X_{it}\beta, -\rho_t)}{\Phi(W_i\gamma)}. \tag{17}$$

Substituting these probabilities into the likelihood function, we arrive at at the following:

$$\begin{aligned}
&\mathcal{L}(\gamma, \beta, \rho, \gamma|\mathbf{Y}, \mathbf{C}, \mathbf{X}, \mathbf{W}) = \\
&\quad \prod_i \Phi(-W_i\gamma)^{1-C_i} \left[\Phi(W_i\gamma, -X_{i1}\beta, -\rho_0)^{1-Y_{i1}} \Phi(W_i\gamma, X_{i1}\beta, \rho_0)^{Y_{i1}}\right]^{C_i} \\
&\quad \times \left[\prod_{t=2}^{T_i} \left(\frac{\Phi(W_i\gamma, -X_{it}\beta, -\rho_t)}{\Phi(W_i\gamma)}\right)^{1-Y_{it}} \left(\frac{\Phi(W_i\gamma, X_{it}\beta, \rho_t)}{\Phi(W_i\gamma)}\right)^{Y_{it}}\right]^{C_i}.
\end{aligned} \tag{18}$$

Note that the value of W_i that obtains when selection occurs is held constant throughout the entire ensuing spell. Even if one has data on how W changes over time, the likelihood requires that one only use the value from the moment of selection since it provides the necessary information about the probability of selection.

5 Monte Carlo Analysis

In this section, we examine the performance of our estimator relative to a discrete event history model through Monte Carlo simulation. This allows us to evaluate its performance relative to a common alternative for a particular set of parameter values.

The data are generated such that in the first period we observe independent variables, X_{i1} and W_i, for each of 1000 individuals, generated according to the following multivariate normal distribution:

$$\begin{pmatrix} X_{i1} \\ W_i \end{pmatrix} \sim MVN\left(\begin{pmatrix} -0.5 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}\right) \tag{19}$$

We hold the values of W fixed for the duration of the spell (i.e., at the values that obtain at the time of selection), which is right-censored after twenty periods, but the values of X change over time to allow for time-varying covariates in the duration equation according to the following formula:

$$X_{it} = X_{it-1} + 0.25 + \nu_{it},$$

for $t > 1$ and with $\nu_{it} \sim N(0, 0.1^2)$. Among the selected observations, the correlation between the two independent variable tends to be a bit lower than this, closer to 0.6 or 0.5 depending on the value of ρ_0.

Using these data, we then generate a variable, C_i, that indicates whether an individual selects into a duration spell.

$$C_i = \begin{cases} 1 & \text{if } -0.5 + 1 \times W_i + \eta_i > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Given these parameter values, about thirty-five percent of the individuals select into the duration process.

Finally, we generate the discrete duration outcome, Y_{it}, according to the following equations:

$$Y^*_{it} = -1 + 0.5 \times X_{it} + \epsilon_{it} > 0, \tag{21}$$

$$Y_{it} = \begin{cases} 1 & \text{if } Y^*_{it} > 0 \text{ and } C_i = 1, \\ 0 & \text{if } Y^*_{it} \leq 0 \text{ and } C_i = 1, \\ . & \text{if } C_i = 0 \text{ or } Y_{it-1} \neq 0. \end{cases} \tag{22}$$

Note that we assume a single failure per selection event; once observations fail in the duration process they exit the risk set. Because the values of X_{it} trend larger over time, the failure rate does not drop too much since observations with larger values tend to fail sooner. While the failure rate depends on the correlation of the error terms, this setup leads to a failure rate around twenty-two percent when ρ_0 is zero; the rate per year either decreases or increases over time depending on the sign of ρ_0. Combined with the initial selection of 350 observations into the duration process, we end up with about 1200 to 1800 individual-year observations in the duration equation, depending on the specific value of the correlation.

Finally, we must parameterize the correlation structure over time in order to introduce non-random sample selection into the duration process. Since the correlation at time t is parameterized as $\rho_t = \rho_0 \exp(-\delta(t-1))$, we must specify values for both the initial correlation, ρ_0 and the decay parameter, δ. Because the value of the decay parameter, in combination with the maximal duration length of twenty years, bounds the maximal correlation, we run our simulations for three different values: 0, 0.3, and 0.4. We also vary the initial correlation from -0.75 to 0.75 by increments of 0.25. Note that $\rho_0 = 0$ corresponds to a situation with no non-random sample selection. For each combination of the two parameters, we check whether the maximal correlation is exceeded, then omit invalid combinations from our simulations. For example, when $\delta = 0.1$, the absolute value of the correlation can not exceed 0.3.

The Monte Carlo simulations are performed by holding the values of the independent variables constant for the entire simulation, drawing new values of the error terms for each trial, calculating the values of C_i and Y_{it}, and estimating two models: our FIML likelihood in Equation 18 and a naive probit discrete time duration model on the observed sample of spells. In both models we cluster the standard errors on individuals. In total, we ran 500 trials for each set of values of the error correlation and decay parameters.[6]

[Figure 3 Here.]

Figure 3 summarizes the results for the parameter of most interest, the slope coefficient, through kernel density plots of the estimates for different correlation and decay parameters. The darker kernel density plot represents the results for the FIML estimator while the lighter plot represents the naive probit results. The decay parameter varies across columns while the correlation changes across rows. The vertical line indicates the true parameter value of 0.5. This figure shows that the FIML estimates suffer from little to no bias whereas the probit estimates exhibit a clear bias — up to twenty-five percent — when there is non-zero correlation. Further, there appears to be little difference in the variability of the two estimates.

[Table 1 Here.]

Tables 1 and 2 report the results of our simulations in more detail and for all parameter values. Table 1 focuses on the parameters from the equation of interest whereas Table 2 focuses on the FIML estimator's additional parameters and the selection equation. The top two panels in Table 1 present the estimates of the slope coefficient, β, while the bottom two panels compare those for the intercept term. Consistent with the plots just reviewed, the results provide evidence in favor of our FIML estimator. The naive probit model always produces estimates further from the true value, with the over- or under-estimation of the slope coefficient approaching twenty-five percent. Note that the bias increases with the correlation for a given rate of decay. In addition, it also increases with slower rates of decay for a given amount of correlation, since the correlation remains larger over the course of the duration. There does seem to be a cost in the form of a slightly larger sampling standard deviation of the estimates, which is not surprising given the greater complexity of the estimator. But when we combine bias and variance through the root mean squared error, the

---

[6] In these simulations we did not restrict the maximal correlation given the estimated decay rate, but estimates rarely violated this condition. For most combinations of these two parameters, over 95% of our estimates satisfied the condition. With larger correlations this dropped to around 80%, but our Monte Carlo results are quite similar to those reported if we exclude these cases. Of course, this restriction has no effect on the bias of the naive estimates to which we compare our estimator.

FIML model outperforms the naive model for both parameters whenever the correlation is not zero. The results for the intercept show similar levels of bias and the FIML model is again preferred by RMSE criterion for all non-zero levels of correlation.

[Table 2 Here.]

Table 2 presents the results for the selection equation parameters and the correlation and decay parameters. The FIML estimator provides accurate estimates of the former for all combinations of correlation and decay parameters. It also provides accurate estimates for the correlation parameter. It does not perform as well for the decay parameter, however, with large deviations from the true value. These deviations are relatively small compared to the standard deviations, however, particularly for smaller true values. Given that these values correspond to very fast decay in the correlation, it is not surprising that these estimates are less precise. In light of the accuracy of the estimates for the other parameters, however, these findings do not appear to undermine the value of our proposed estimator.

6 The Duration of Exchange Rate Regimes

In this section we attempt to explain the length of pegged exchange rate regimes with the selection corrected duration estimator developed above. Although no paper we are aware of in the literature on peg duration has investigated selection effects, selection on unobservables may be an important problem.

Obstfeld and Rogoff (1995) argued that ‘credibility’ was the key to maintaining a peg, that credibility was becoming increasingly hard to earn in the face of global capital flows, and that it was “hard to quantify” the political reasons some nations had more success than others in maintaining pegs. Many countries during the 1990s abandoned pegged rates of intermediate “hardness” in the midst of financial turbulence and financial crises. The political costs of defending pegs were seemingly too high, and countries found it increasingly difficult to generate the credibility needed to maintain a pegged exchange rate (Obstfeld and Rogoff 1995). Only a limited group of economies were willing to subordinate monetary autonomy to the defense of a fixed exchange rate regime.

The prognosis was grim for the ability of countries to maintain fixed exchange rate regimes as of the mid-1990s, and many economists took the view that intermediate pegs were increasingly short-lived as global capital markets burgeoned. However a handful of exceptional countries had managed to maintain pegged exchange rate regimes for longer than five years. For these countries, Obstfeld and Rogoff noted that one political factor "though difficult to quantify, is that all potential ruling groups...share a strong consensus on the primacy of the fixed-rate commitment" (Obstfeld and Rogoff 1995, pp. 87-88).

Klein and Shambaugh (2008) recently offered a revisionist view of peg duration. They accept that the data show that most pegs break after a short period, but they also note that a significant proportion of peg spells (30 percent) have in fact lasted longer than five years. Klein and Shambaugh estimate a hazard model for exchange rate peg spells controlling for duration dependence and the length of the preceding float spell but not for other economic or political fundamentals. They find evidence for positive duration dependence–the longer a peg lasts, the less likely it is to break. Also, longer periods of floating, prior to a peg spell, are associated with shorter pegs and short preceding floats are associated with longer peg spells.

Our reading of the this and other related literature on peg duration is that political factors and unobservable, hard to quantify factors related to 'credibility' are at play in determining the longevity of a peg spell. A couple of thought experiments can illustrate how selection could matter.

First, there may be a relationship between the (perceived) ability and/or willingness to sustain an exchange rate peg and the choice of whether to opt into a fixed exchange rate in the first place. If so, then empirical work on the duration of pegs should account for selection effects before making inferences about the determinants of duration. Certain types of countries that have an unobservable or hard to measure ability to sustain a peg, or those which expect to gain the most, or lose the least, from a peg are the types of countries that might select a fixed exchange rate in the first place.

Alternatively, weaker countries might select into pegs. It is possible that short-run political considerations rather than solid fundamentals tip countries into pegs. Exchange rate based stabilizations are often viewed (or recommended) as quick and effective means of eliminating volatility. In the medium term it is possible that other political goals and concerns trump original policy. Also, policy makers may systematically underestimate any of the following: the ability of capital markets to terminate exchange rate pegs, the ability to deal with such an attack, or the ability to maintain macroeconomic policy consistent with the peg commitment. Selection is operative in all of these cases. Pegs would be more likely to fail again for reasons that are hard to quantify.

Surprise changes to the environment are a component of the discussion in Obstfeld and Rogoff. There they seem to suggest that it was increasingly difficult to establish sufficient credibility to maintain a peg as of the mid-1990s in the midst of rising cross-border capital flows. Perhaps the series of spectacular currency crashes in the 1990s was a learning experience for countries since they may not have anticipated the ease with which international capital markets could put them to the test via speculative attacks. Fixed exchange rates were the conventional policy prescription for most countries post-Bretton Woods, through the EMS stage of European Monetary Union and even during the early years of the ‘Washington Consensus’ as a means for stabilization, but policy makers were slow to realize the disruptive capacity of global capital markets until after 1997-98.

Quite obviously, the durability of an exchange rate peg depends in large part on the policy, preferences and the political capabilities of countries to successively maintain their peg from year-to-year. Forward looking expectations and perceptions by politicians and economic actors that influence policy and unobservable but related factors could influence the decision to join or not in the first place.

Sturzenegger, Levy-Yeyati and Reggio (2007) find that several political factors are important in explaining why countries adopt fixed exchange rates. Their evidence supports the idea that stable and strong governments are more likely to adopt a peg since they will be able to take actions consistent with a peg even if this implies eliminating a deficit for example. Also they show that a government with higher numbers of veto players are less likely to adopt a peg. Such divisions in the policy making process could make adopting a peg more difficult, but they could also be related to the ability to appropriately adjust in the face of a shock if a peg were to be adopted.

In the political science literature, a significant amount of research has focused on credibility and political factors in explaining the demise of currency pegs. Leblang and Satyanath (2008) and Leblang and Bernhard (2000) find evidence that political instability and political uncertainty are significant determinants of currency crises. Leblang and Satyanath argue that this type of result is consistent with a model where a speculative attack is more likely when agents have a wider range of beliefs about government policy. Leblang and Satyanath (2008) also examine the idea that divided governments characterized by uncertain preferences and delays in decision making are more likely to have high costs of responding to shocks. Unified governments in their samples are found to be less likely to have a currency crisis since they can respond to shocks with greater resolve.

We analyze these issues in terms of both exchange rate regime choice and duration while also controlling for the possibility that these factors are correlated with unobservables. Such an empirical strategy relates directly to the discussion above. In the context of the political theories of currency crises discussed above, it could be the case that nations size up the expected benefits and costs of joining a peg which include the possibility of a currency crisis when making the exchange rate regime choice. These costs and benefits would depend on the expected duration of the spell (i.e., the likelihood of a currency crisis or changes in policy preferences in the future due to economic of political change). And these in turn could be related to the political and economic characteristics of a country at the time of choosing to opt into a pegged exchange rate regime.

For instance, assume politicians want to avoid the economic disruption and political fallout that currency crises entail or disdain the idea of a major policy flip-flop in the proceeding years. Then it would be expected that only politicians or governments that view themselves as capable and willing to take the necessary actions to defend a peg would opt in. Alternatively, politicians that have (unobservable) short time horizons or who need a rapid exchange rate based-stabilization for short-term political gain may be more likely to opt in but also they may more easily fall out of their pegs.

6.1 Methods and Data

The variable to be explained is the duration of a pegged exchange rate spell, and we apply our estimator to do so. Our data include information on whether countries establish a peg and how long that peg is maintained. Because countries can start their pegs in any year, our selection equation is itself a duration model. This involves only a minor extension to our likelihood since a discrete duration model can be estimated with any appropriate discrete choice model. To adapt our estimator, one only needs to subscript the selection equation variables by time.[7]

---

[7] Crucially, one must use the values of the independent variables from the selection equation that obtain at the time of selection when calculating the likelihood for the entire spell. Even if these variables change over time, it is their values at the moment of selection that provides information about the unobserved components at the time of selection.

We use the Klein and Shambaugh de facto classification for peg spells to analyze the issue of duration.[8] Klein and Shambaugh's measure classifies a country as having a peg during a calendar year if the end of month exchange rate stayed within a band of $\pm 2\%$ against another reference currency in each month of a calendar year and over the course of that year. They argue that the assignment of countries to pegs is robust to the choice of bandwidth. This data set is unlike the well known Reinhart and Rogoff (2004) de facto classification because in the Klein and Shambaugh data parity changes mean a peg spell has ended. Reinhart and Rogoff smooth their data so that one-time parity changes do not end a spell. In this way, the variable of analysis for Reinhart and Rogoff is smoothed exchange rate policy rather than any particular exchange rate.

Our data cover 1973 to 2000 and include pegged exchange rate spells that begin after 1972. During this time period, we have 334 instances of pegged exchange rate spells from 125 different countries. These regimes last an average of 3.7 years with a median duration of one year, including sixty-four ongoing spells in 2000.

The selection model includes openness to trade, the logarithm of GDP, whether the country has a large inflation in the recent past, whether a country had restrictions on the capital account, and the number of years since 1973.[9] The economic determinants we include in the duration model are: GDP growth, trade openness, the trade deficit (data from the IFS–International Financial Statistics), international reserves relative to imports (IFS data), and the time in years since the spell began.[10]

For the political determinants of regime duration we include a measure of political stability and a measure of divided government. The political stability variable indicates the amount of recent turnover in the government from the Database of Political Institutions (Beck et al. 2003). This variable measures the extent of turnover in the key decision makers of a government in any

---

[8] It has become well known in the 1990s that countries' actual exchange rate policies differ from what they report to the IMF or announce to the public. Since what matters is not what policy makers say but what actually happens, many authors have now turned to looking at what actually happened to the exchange rate via such de facto classifications.

[9] Openness (i.e., exports plus imports divided by GDP and PPP) and PPP-adjusted GDP come from the Penn World Tables. The large inflation indicator is 1 if a country had a freely falling exchange rate, which would typically be associated with high inflation, as defined in Reinhart and Rogoff between the current year and 1950. The measure of capital account openness is from Chinn and Ito (2006).

[10] Many of the variables that we wanted to include were missing in almost half the cases. This problem is exacerbated here since when a variable that is included in the selection equation is missing in the year a duration begins, the entire duration is omitted. The variables we were forced to omit include trade balance, reserves, capital controls, openness to capital flows, and political changes.

one year. We also include a variable from the same data set that measures the extent to which the executive controls the legislature. This divided government measure is equal to 1 if the chief executive's party is in control of the legislature and 0 otherwise.[11] Both variables are lagged by one year to avoid simultaneity issues.

6.2 Results

Table 3 presents the results from three different models: the first provides estimates from our duration with selection estimator while the latter two present separate models of the decision to start a peg and the duration of observed pegs. Consider first the results from the former. For the selection equation we find that size and past experience matter. Smaller countries are more likely to adopt a peg although the coefficient on the log of GDP barely misses weak significance ($p = .103$). Also, countries that had high inflation or moved from a peg to a "freely falling peg" in the past are significantly less likely to choose a peg. Other variables are not statistically significant. For the duration model the only statistically significant variable is political instability. A large turnover in the previous year is associated with a higher likelihood of an exit from a peg with a p-value of .036. The fact that other variables are not statistically significant does not mean there is not other information available: a comparison of the probit duration model that does not control for selection and the probit model that does reveals some interesting information.

[Table 3 Here.]

Importantly, our results indicate that accounting for possible non-random sample selection matters for understanding the duration of exchange rate pegs. These differences manifest themselves in a number of ways. First, the estimate of the correlation between the selection model and the duration model is positive and significant at the .05 level.[12] The parameter value indicates a correlation of 0.81 between the errors for the equation modeling the decision to start a peg and the error for first year of the duration of a new spell. The significant decay parameter indicates that this correlation decreases over the course of that spell. It decreases by about half each year, dropping to 0.46 in the second year and 0.26 in the third year. Further, application of Equation 7 indicates that

---

[11] These are the variables labeled STABS and ALLHOUSE, respectively, and are the same variables used in Leblang and Satyanath (2008).

[12] In order to facilitate estimation we use the inverse of Fischer's Z transformation so that ρ_0 lies between -1 and 1 and an exponential transformation so that δ is positive.

this combination of parameter values corresponds to a permissible covariance matrix (the longest observed spell is 26 years). Note that of the 164 pegs in our estimation sample, 98 end in the their first year and that almost 143 have ended by their third year.[13] Thus the selection effect has a strong influence on almost every peg since it is largest exactly when most pegs are ending.

The positive correlation between the error term in the duration model and the selection equation supports the idea that countries implement pegs that, though they may be difficult to maintain, offer immediate short-term gains that offset the higher probability of failure over time. It is also evidence against the idea that only countries that are stronger or more capable for unobservable reasons choose to implement pegs.

Accounting for selection changes the interpretation of the effect of a number of variables on the duration of pegged exchange rates. First, the coefficient for the reserves ratio is negative and significant at the .10 level in the naive probit model ($p = .052$) but is not near significance in the selection model ($p = 0.24$). Note that the change in significance results mainly from a reduction in the magnitude of the coefficient rather than an inflation in the standard error. Accounting for non-random sample selection appears to eliminate induced correlation between reserves and unobserved factors. Second, the opposite occurs for political instability, which becomes significant at the .05 level once we account for selection (the p-value goes from 0.141 to $= .036$). This is compatible with the idea that political factors matter more than the economic ability to maintain a peg with reserve backing. Without controlling for selection, it appears that the impact of economic variables is overstated.

Third, after accounting for selection there remains no significant duration dependence. Although the naive probit indicates negative duration dependence, it appears that this result is almost entirely driven by unobservables rather than duration dependence per se. Fourth, a direct comparison of the two models also provides evidence for our estimator. A likelihood ratio test comparing the combined likelihoods of the two independent models to the likelihood of the combined model produces a χ^2_2 test statistic value of 24.3, which has a p value less than .001.[14]

[Figure 4 Here.]

---

[13] We lose the other 169 pegs in our data set due to missing data, which is particularly problematic in this context. If a covariate explaining selection is missing at the time of entry, the entire spell is lost since we need to include its value when calculating the likelihood contribution for each subsequent year of the associated spell.

[14] The formula for the test statistic is $-2((-213.092 - 497.228) - (-698.165))$.

In order to better illustrate the differences in the findings, Figure 4 plots the hazard function for exchange rate pegs from the naive and selection models over time. We hold all independent variables fixed at their median values, with the exception of time, which we increase from the first year of a spell up to fifteen years. Because the naive probit model exhibits negative duration dependence, the hazard steadily decreases over time. The duration model with selection, however, exhibits two competing forces with different effects: one from the insignificant, but positive, effect of duration dependence, the other from the decay of the positive correlation over time. In order to better distinguish these two forces, we predicted the hazard both accounting for and ignoring the effect of duration dependence. The lighter dashed line isolates the effect of the correlation as it decays over time and shows how the hazard decreases in response. After about eight years, the correlation is basically zero and the hazard remains constant over time. The black dashed line then incorporates the estimate of duration dependence, which, while modest at first, ultimately pushes the hazard to start increasing after about six years. Given that most of the observed pegs in our sample end in the first few years, this figure shows that the naive model tends to understate the chance that a peg will fail in the first year by almost 0.1, but then overstates it for the next few years by about the same amount.

7 Discussion and Conclusions

In this paper we discuss the problem of extending existing estimators for duration models that allow for non-random selection to cases with time-varying covariates. This extension is not straightforward and, we believe, is best handled by moving from continuous-time duration models like the Weibull to discrete-time duration models. Here we use the bivariate normal distribution to construct a full-information maximum likelihood estimator, essentially combining standard models for stochastic truncation and stochastic censoring with binary dependent variables of interest. Monte Carlo simulation shows that our estimator generally outperforms a naive discrete time duration estimator when the correlation is nontrivial. An empirical application of this estimator to the duration of pegged exchange rate regimes shows that the presence of non-random sample selection greatly influences our conclusions about which factors matter.

As demonstrated in our empirical application, it is straightforward to extend our estimator to allow for a discrete time duration process in the selection equation. An individual has the chance to begin the duration process at different points in time, though all information about selection occurs in the period in which they ultimately enter the duration process. Many political science applications of this sort raise an additional issue: repeated events. For example, countries may select into conflicts, see the conflicts end, but then select into new conflicts in the future. Our estimator would allow for this and many of the covariate-based adjustments for repeated events (see, e.g., Box-Steffensmeier and Zorn 2002).

Repeated events suggest the possibility not only of repeated selection into events, but also from events back to the selection process. That is, countries may select into spells of conflict, but when they end them they also select back into spells of peace. Accounting for this would greatly complicate our estimator. If the initial spell of conflict must condition on factors at the time of selection, then the subsequent spell of peace would have to condition on factors at the end of conflict and factors at the time of selection into that conflict. Even this simple case would involve a trivariate distribution, with the dimensions continuing to grow with the number of spells. While the believe this would be a useful extension to develop, at this point we believe it would be difficult to get estimates in most practical situations.

We also want to consider alternate approaches that treat the unobserved heterogeneity through correlated random effects. This approach has been used to generate panel selection estimators for continuous outcomes of interest (e.g., Kyriazidou 1997, Vella 1998). One might be able to extend these estimators to allow for discrete outcomes of interest. One important difference involves the timing of the selection mechanism, however: panel selection models assume that the selection process occurs each period whereas our estimator assumes that it occurs at the beginning of a spell. Certainly one could think of empirical applications in which it would be best to model the selection process as re-occurring in each period of an ongoing spell, so this approach may prove valuable on its own. Extending it to allow for a single selection decision for each spell would correspond more closely to the structure of out estimator and, we believe, would apply to a wider variety of empirical situations in political science.

References

Beardsley, Kyle and Victor Asal. 2009. “Winning with the Bomb.” *Journal of Conflict Resolution* 53 (2): 278-301.

Beck, Nathaniel, Jonathan N. Katz and Richard Tucker. 1998. “Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable.” *American Journal of Political Science* 42(4): 1260-1288.

Beck, T.; G. Clarke; A. Groff; P. Keefer; and P. Walsh. 2003. *Database of Political Institutions.* Washington DC: World Bank.

Boehmke, Frederick J., Daniel Morey and Megan Shannon. 2006. “Selection Bias and Continuous-Time Duration Models: Consequences and a Proposed Solution.” *American Journal of Political Science* 50: 192-207.

Box-Steffensmeier, Janet M. and Bradford D. Jones. 2004. *Event History Modeling: A Guide for Social Sciences.* Cambridge: Cambridge University Press.

Box-Steffensmeier, Janet M., and Christopher J.W. Zorn. 2002. “Duration Models for Repeated Events.” *The Journal of Politics* 64 (November): 1069-94.

Chinn, Menzie and Hiro Ito. 2006. “What Matters for Financial Development? Capital Controls, Institutions, and Interactions.” *Journal of Development Economics* 61(1): 163-192.

Dubin, Jeffrey A. and Douglas Rivers. 1990. “Selection Bias in Linear Regression, Logit and Probit Models.” *Sociological Methods and Research* 18:354-365.

Heckman, James J. 1976. “The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models.” *Annals of Economic and Social Measurement* 5/4:475-492.

Heckman, James J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47: 153-161.

Klein, Michael W. and Jay C. Shambaugh. 2008. “The Nature of Exchange Rate Regimes.” *Journal of International Economics* 75 (1): 70-92.

Klein, Michael W. and Nancy P. Marion. 1997. “Explaining the Duration of Exchange Rate Pegs.” *Journal of Development Economics* vol. 54 pp. 387-404.

Kyriazidou, E. 1997. “Estimation of a Panel Data Sample Selection Model.” *Econometrica* 65 (6): 1335-1364.

Leblang, David and Shanker Satyanath. 2008. “Politically Generated Uncertainty and Currency Crises: Theory, Tests, and Forecasts.” *Journal of International Money and Finance* 27 (30): 480-497.

Leblang, David and William Bernhard. 2000. “The Politics of Speculative Attacks in Industrial Democracies.” *International Organization* 54 (2): 291-324.

Long, Andrew G.; Timothy Nordstrom; and Kyeonghi Baek. 2007. “Allying for Peace: Treaty Obligations and Conflict between Allies.” *Journal of Politics* 69 (4): 1103-1117.

Obstfeld, Maurice, and Kenneth Rogoff. 1995. “The Mirage of Fixed Exchange Rates.” *Journal of Economic Perspectives* 9 (Fall): 73-96.

Pang, Xun. 2008. “Binary and Ordinal Time Series with AR(p) Errors: Bayesian Model Determination for Latent High-Order Markov Process.” Paper presented at the 25th annual summer meeting of the Society for Political Methodology.

Reinhart, Carmen and Kenneth Rogoff. 2004. “The Modern History of Exchange Arrangements: A Reinterpretation.” *Quarterly Journal of Economics* Vol. CXIX, No. 1, pp. 1-48

Maddala, G.G. 1983. *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

Mooney, Christopher. 1997. *Monte Carlo Simulation.* Sage Publications.

Prieger, James E. 2002. “A Flexible Parametric Selection Model for Non-Normal Data with Application to Health Care Usage.” *Journal of Applied Econometrics* 17: 367-392.

Smith, Murray D. 2003. “Modeling Sample Selection Using Archimedean Copulas.” *Econometrics Journal* 6: 99-123.

Sturzenegger, Federico; Eduardo Levy-Yeyati; and Iliana Reggio. 2007. “On the Endogeneity of Exchange Rate Regimes.” mimeo. Kennedy School of Government, Harvard University.

Vella, Francis. 1998. “Estimating Models with Sample Selection Bias: A Survey.” *The Journal of Human Resources* 33(1): 127-169.

Table 1: Monte Carlo Simulation Results Comparing FIML Estimator to Probit Discrete Duration Estimator, Varying the Decay Parameter and the Error Correlation

		Error Correlation													
Decay		-0.75	-0.5	-0.25	0.0	0.25	0.5	0.75	-0.75	-0.5	-0.25	0.0	0.25	0.5	0.75
Parameter		*FIML Duration (Slope)*							*Naive Probit Duration (Slope)*						
0.1	Mean			0.518	0.502	0.510					0.575	0.502	0.457		
	SD			0.057	0.052	0.059					0.044	0.043	0.046		
	RMSE			0.060	0.052	0.060					0.087	0.043	0.063		
0.3	Mean		0.517	0.507	0.498	0.506	0.518			0.628	0.562	0.500	0.453	0.408	
	SD		0.056	0.053	0.054	0.053	0.049			0.044	0.043	0.046	0.043	0.045	
	RMSE		0.059	0.054	0.055	0.054	0.052			0.136	0.076	0.046	0.063	0.102	
1.0	Mean	0.510	0.505	0.496	0.503	0.507	0.507	0.509	0.611	0.580	0.541	0.503	0.464	0.416	0.374
	SD	0.054	0.055	0.054	0.052	0.052	0.048	0.042	0.048	0.047	0.046	0.044	0.046	0.045	0.043
	RMSE	0.055	0.055	0.054	0.052	0.052	0.049	0.043	0.121	0.093	0.062	0.044	0.059	0.095	0.133
		FIML Duration (Intercept)							*Naive Probit Duration (Intercept)*						
0.1	Mean			-1.016	-0.997	-0.976					-1.187	-0.999	-0.819		
	SD			0.118	0.093	0.120					0.048	0.044	0.046		
	RMSE			0.529	0.506	0.491					0.688	0.501	0.322		
0.3	Mean		-0.995	-0.996	-0.994	-0.998	-1.004			-1.261	-1.131	-1.000	-0.871	-0.729	
	SD		0.108	0.097	0.092	0.096	0.077			0.047	0.045	0.045	0.045	0.044	
	RMSE		0.507	0.506	0.503	0.507	0.510			0.762	0.632	0.502	0.373	0.233	
1.0	Mean	-0.995	-0.993	-0.980	-1.004	-1.010	-1.018	-1.010	-1.208	-1.146	-1.076	-1.003	-0.923	-0.840	-0.747
	SD	0.077	0.083	0.081	0.094	0.079	0.072	0.060	0.048	0.049	0.046	0.045	0.046	0.046	0.050
	RMSE	0.501	0.500	0.487	0.513	0.517	0.523	0.514	0.710	0.648	0.577	0.505	0.425	0.344	0.252

Notes. Results based on 100 trials, with 1000 units and spells lasting up to 20 periods. Estimates from a handful of trials that failed to converge are excluded. Holes represent infeasible combinations of correlation and decay parameters. See paper for additional details.

Table 2: Monte Carlo Simulation Results for Selection Equation and Selection Parameters, Varying the Decay Parameter and the Error Correlation

		Error Correlation													
Decay Parameter		-0.75	-0.5	-0.25	0.0	0.25	0.5	0.75	-0.75	-0.5	-0.25	0.0	0.25	0.5	0.75
		Selection Equation (Slope)							*Selection Equation (Intercept)*						
0.1	Mean			1.009	1.004	1.002					-0.502	-0.502	-0.497		
	SD			0.064	0.062	0.065					0.052	0.047	0.049		
	RMSE			0.064	0.062	0.065					0.501	0.500	0.505		
0.3	Mean		1.003	1.003	1.007	1.003	1.005			-0.501	-0.504	-0.500	-0.501	-0.504	
	SD		0.065	0.064	0.066	0.061	0.059			0.052	0.051	0.050	0.051	0.048	
	RMSE		0.065	0.064	0.066	0.061	0.059			0.501	0.499	0.503	0.502	0.499	
1.0	Mean	1.002	1.000	1.006	1.002	1.003	1.002	1.000	-0.504	-0.500	-0.500	-0.502	-0.503	-0.501	-0.500
	SD	0.060	0.063	0.060	0.065	0.062	0.062	0.060	0.046	0.051	0.051	0.047	0.048	0.049	0.051
	RMSE	0.060	0.063	0.060	0.065	0.062	0.062	0.060	0.498	0.502	0.502	0.500	0.499	0.501	0.502
		Correlation Parameter							*Correlation Decay Parameter*						
0.1	Mean			-0.251	0.005	0.243					-1.785	-1.249	-3.383		
	SD			0.155	0.166	0.194					3.368	5.761	6.656		
	RMSE			0.155	0.166	0.194					3.407	5.856	6.744		
0.3	Mean		-0.514	-0.266	-0.007	0.259	0.528			-1.025	-0.934	-1.026	-1.395	-1.494	
	SD		0.141	0.140	0.169	0.174	0.206			0.656	2.354	5.446	3.919	0.694	
	RMSE		0.142	0.141	0.169	0.174	0.208			0.679	2.369	5.449	3.924	0.754	
1.0	Mean	-0.762	-0.514	-0.283	0.000	0.258	0.529	0.767	0.031	0.168	0.194	-0.922	-0.263	-0.117	-0.285
	SD	0.420	0.138	0.127	0.163	0.154	0.276	0.833	1.096	1.295	2.035	4.705	3.120	0.802	0.428
	RMSE	0.420	0.138	0.131	0.163	0.154	0.278	0.834	1.096	1.306	2.045	4.794	3.131	0.811	0.514

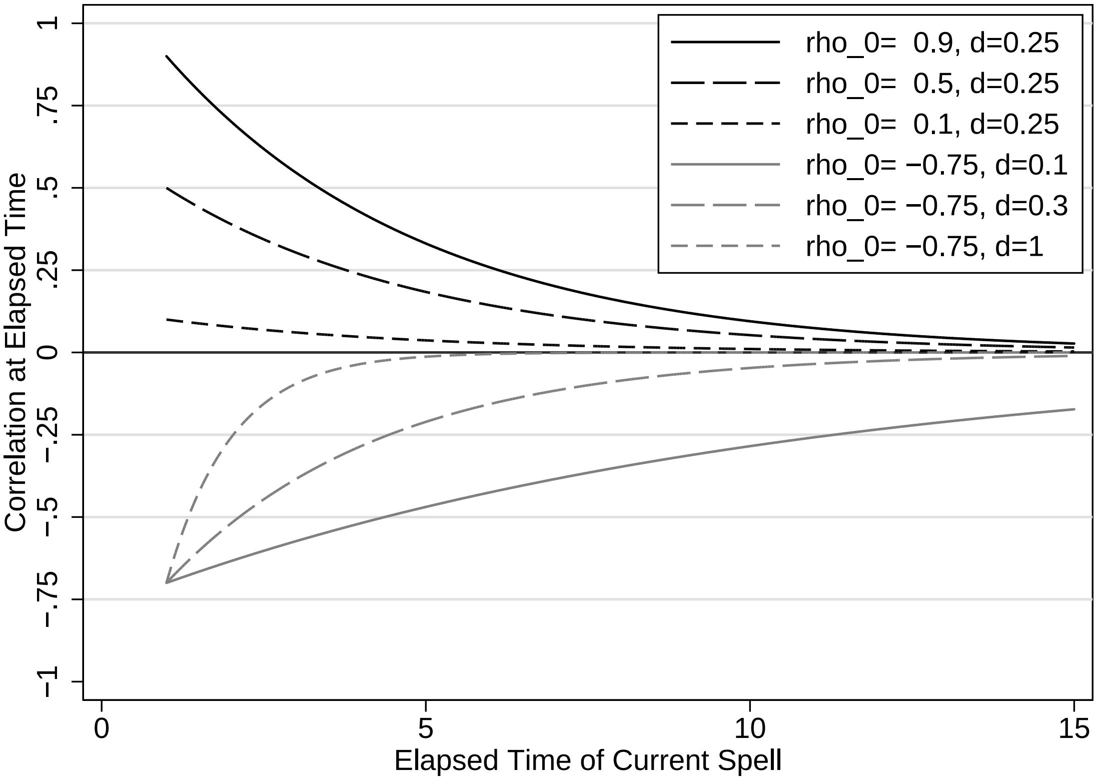
Notes. Results based on 100 trials, with 1000 units and spells lasting up to 20 periods. Estimates from a handful of trials that failed to converge are excluded. Holes represent infeasible combinations of correlation and decay parameters. See paper for additional details.

Table 3: FIML Duration and Selection Estimator versus Naive Probit Estimates of Pegged Exchange Rate Regime Initiation and Duration, 1960-2004

	FIML Duration and Selection		**Naive Probit**			
			Duration		**Selection**	
	Coef.	SE	Coef.	SE	Coef.	SE
Selection						
Trade Openness	0.028	0.279			0.036	0.285
GDP (log)	-0.048	0.029			-0.047	0.031
Recent Hyperinflation	$-0.383**$	0.115			$-0.430**$	0.127
Unified Government (Lagged)	0.080	0.115			0.063	0.117
Stability (Lagged)	0.106	0.153			0.129	0.153
Capital Account Openness	-0.004	0.046			-0.016	0.055
Time	0.024	0.033			0.032	0.036
Time Squared	-0.000	0.001			-0.001	0.001
constant	-0.320	0.801			-0.351	0.828
Duration						
Trade Openness	-0.092	0.380	-0.268	0.500		
GDP Growth	0.613	0.504	0.808	0.679		
Reserves/Imports	-0.386	0.331	$-0.722*$	0.372		
Exports/Imports (logged)	-0.045	0.172	0.012	0.256		
Stability (Lagged)	$0.440*$	0.210	0.370	0.251		
Unified Government (Lagged)	-0.074	0.161	-0.175	0.200		
Spell Time	0.035	0.034	$-0.182**$	0.038		
constant	$-1.168**$	0.315	0.250	0.336		
Correlation						
$F^{-1}(\rho)$	$1.151**$	0.248				
ρ	$0.818**$	0.082				
$ln(\delta)$	$-0.548*$	0.239				
δ	$0.578**$	0.138				
Observations	1628		353		1439	
Final Log-likelihood	-698.165		-213.092		-497.228	

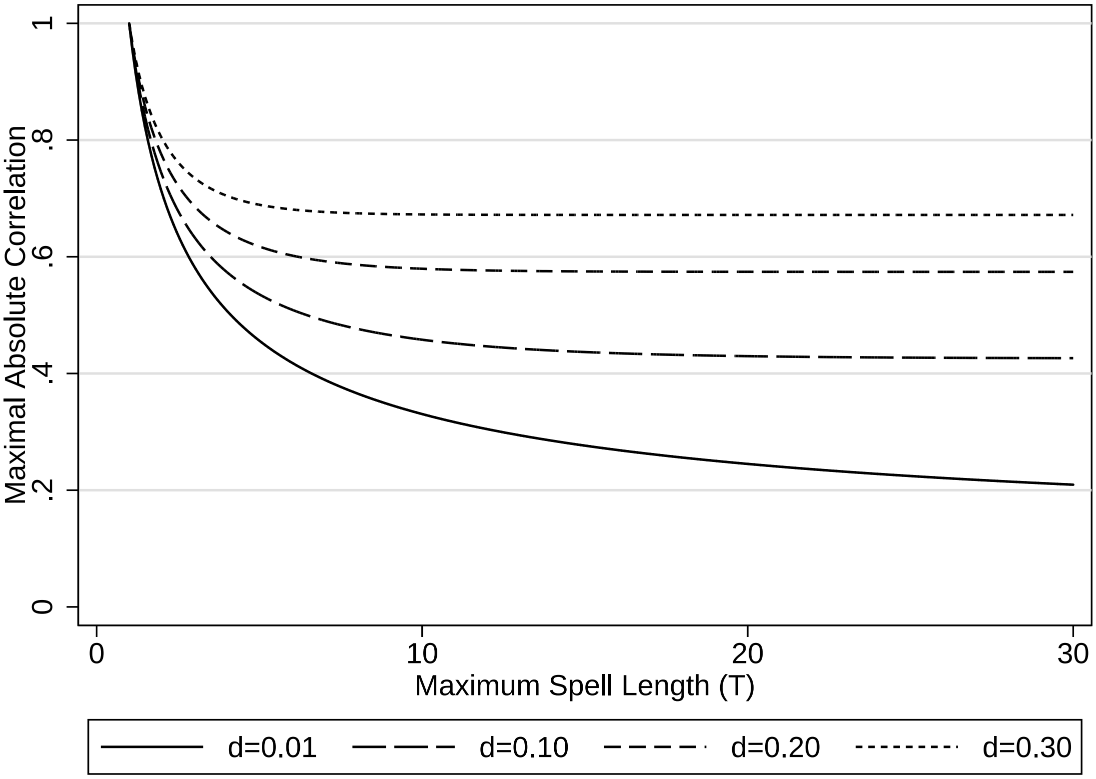
Notes. * indicates $p \leq .10$ with a two-tailed test; ** indicates $p \leq .05$. Standard errors clustered on country. Likelihood ratio test for two independent equations versus constrained FIML estimator: $\chi^2_2 = 24.3$ $(p = .0000053)$.

Figure 1: Examples of Correlation Over Time, Varying Initial Correlation and Decay Parameter



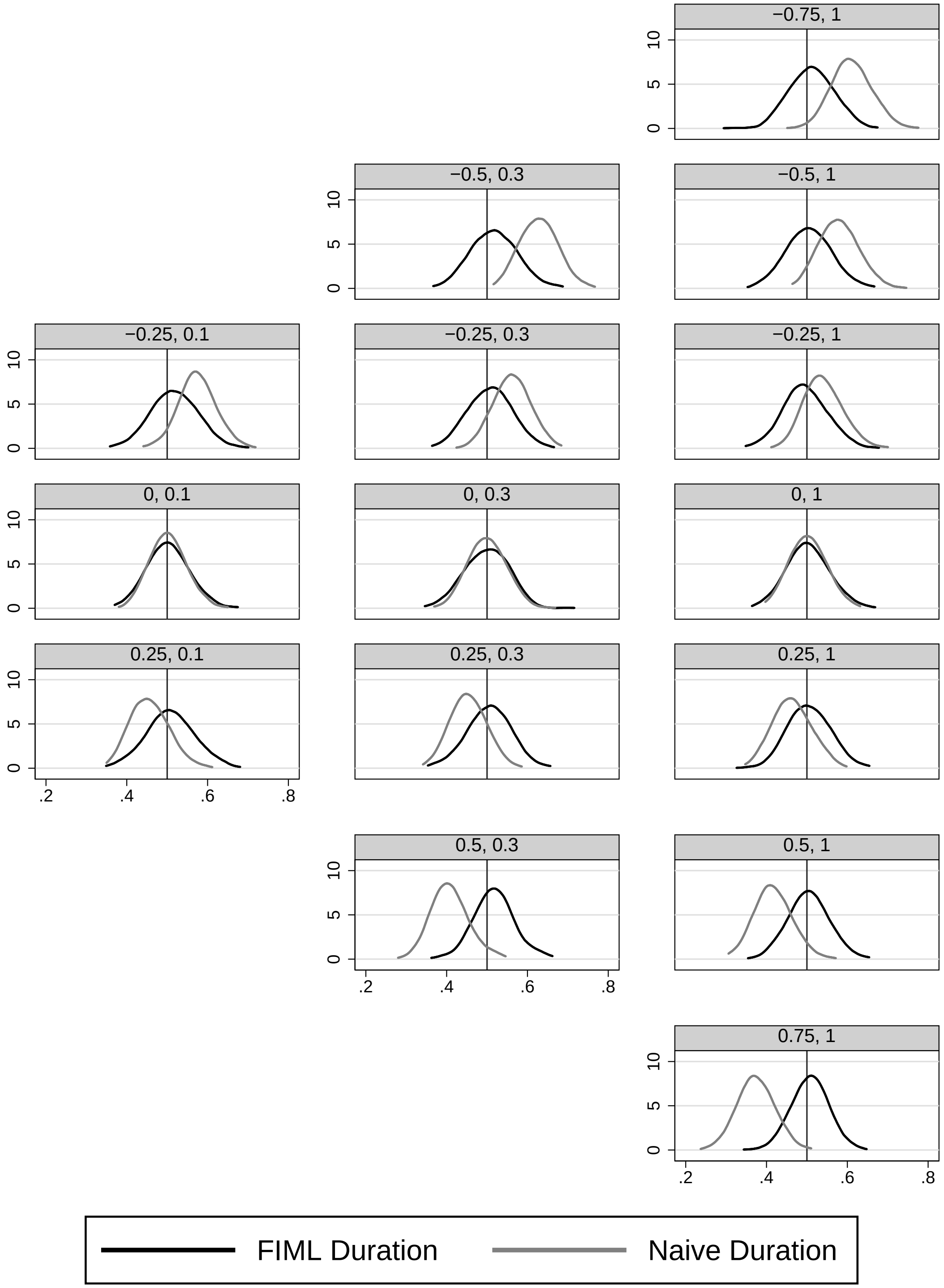
Notes. $\rho_t = \rho_0 \exp(-\delta(t-1))$.

Figure 2: Maximal Absolute Correlation in the First Period of a Spell by Maximum Spell Length, for Different Rates of Decay



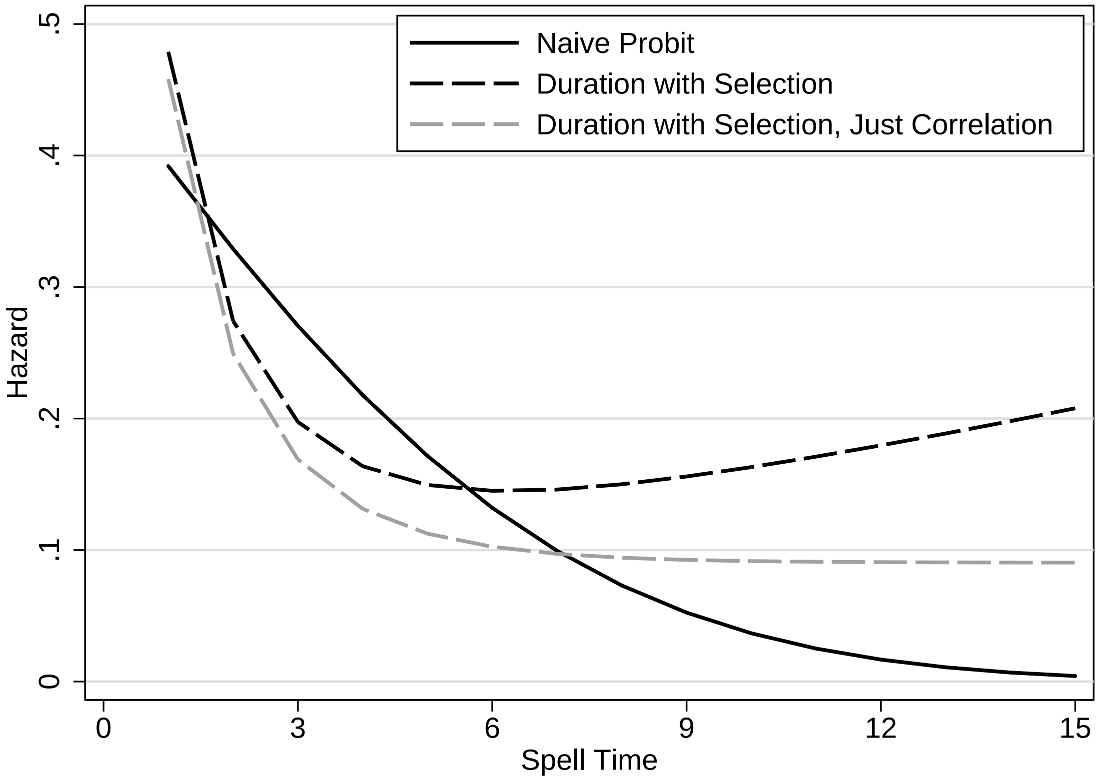
Notes. Constraint given by $|\rho_0| \leq \sqrt{\frac{1-\exp(-2\delta)}{1-\exp(-2\delta T)}}$.

Figure 3: Kernel Density Plots of Slope Coefficient Estimates from Monte Carlo Analysis, Varying Initial Correlation and Correlation Decay Rate



Notes. Constructed from results reported in Table 1. The first parameter listed is the correlation between the errors; the second parameter listed is the correlation decay parameter. Holes represent infeasible combinations of correlation and decay parameters.

Figure 4: Predicted Hazards for Pegged Exchange Rate Regime Durations



Notes. Constructed from results reported in Table 3. All variables held constant at the median values.