The Dynamics of Campaign Contributions in U.S. House Elections

by

Walter R. Mebane, Jr.[†]

and

Jonathan Wand[‡]

August 29, 1999

Prepared for delivery at the 1999 Annual Meeting of the American Political Science Association, September 2–5, 1999, Division 8: Political Methodology Section. A previous version of this paper was presented at the Summer Political Methodology Meeting, 15–18 July 1999, Texas A&M University. Thanks to Robert Biersack for his help in interpreting the FEC contributions data, and to Jonathan Hoffman and Erich Sachse for research assistance. Data were made available in part by the Cornell Institute for Social and Economic Research (CISER) and the Inter-University Consortium for Political and Social Research. Thanks to CISER for letting gaea help p2c2e and macht with the computing. Wand gratefully acknowledges support from the Social Science and Humanities Research Council of Canada (grant 752-98-0374). The authors share equal responsibility for all errors.

[†] Associate Professor, Department of Government, Cornell University, wrm1@macht.arts.cornell.edu.

[‡] Doctoral Candidate, Department of Government, Cornell University, jnw4@cornell.edu.

Abstract

The Dynamics of Campaign Contributions in U.S. House Elections

We use Federal Election Commission itemized contributions data from 1984 to estimate a model of campaign contributions in U.S. House elections. The model is a dynamic system of conditional compound Poisson processes in which there are contributions from both individuals and political action committees (PACs). The model includes random effects to allow for unobserved heterogeneity among districts and candidates. The dynamic effects measure how contributions to one candidate react to contributions to other candidates, as well as how contributions from individuals interact with contributions from PACs. We test the hypothesis that some candidates received higher contributions because of PAC endorsements. We also test whether national expectations about presidential election outcomes affect contributions to House candidates, as predicted by a policy moderating model. We use a Monte Carlo EM algorithm to optimize the likelihood of the model in specifications that include more than one random effect.

Introduction

The large positive correlation between money and electoral outcomes among non-incumbent U.S. House candidates and the recent easy availability of Federal Election Commission (FEC) itemized contributions data has sparked the search for adequate models of how some candidates are able to raise money and others fail. There is increasing attention to the analysis of fundraising and its evolution through an election cycle (Biersack, Herrnson and Wilcox 1993, Krasno, Green and Cowden 1994, Box-Steffensmeier 1996, Himmelberg and Wawro 1998, Wand and Mebane 1999). To estimate a model of contributions that allows for contributions from multiple sources, such as both individuals and political action committees (PACs), allows for unobserved heterogeneity and specifies a plausible data generating process, we adopt the framework of generalized linear mixed models (GLMMs). Our model is a system of conditional compound Poisson processes that have means that are functions of fixed effects and normally distributed random effects: we assume that conditional on the values of unobserved random effects, there is a Poisson process for the number of contributions candidates receive, and given that a contribution occurred, there is a separate process that determines the dollar amount of each contribution. The model includes specification of a stochastic dynamic system of contributions, by including effects of lagged and cross-lagged contributions amounts in the model of the number of contributions. To estimate the model we use a Monte Carlo EM (MCEM) algorithm similar to the one introduced by McCulloch (1997).

In the remaining part of this introduction we outline some of the substantive issues pertaining to contributions that we address in the current analysis. In the following sections we outline the motivation for modeling contributions as a count process and present the statistical model. Then we present a detailed account of the models we estimate and the results.

Fundraising is a dynamic process, with a sequence of opportunities for a candidate to appeal to potential contributors during an election campaign. In particular, acquiring money from one type of contributor may affect a candidate's ability to raise money from other sources. Many political action committees (PACs) are known to use a candidate's past success in raising individual contributions in deciding whether to contribute (Sabato 1984, Biersack, Herrnson and Wilcox 1994). Similarly, the actions of one PAC may affect the behavior of other PACs, as well as of individual contributors: an endorsement by a PAC may help coordinate targeted contributions to particular candidates. Even if individuals and PACs are not wittingly conditioning their contributions on the past behavior of others, money from any source can pay for direct mailing lists, professional campaign staff, and other instruments that facilitate raising money from different sources. In addition, successful fundraising efforts by a candidate's opponent may also encourage greater effort by a candidate and provide extra leverage in requesting additional donations from the candidate's supporters. An adequate model therefore needs to allow for a system of processes for contributions from different sources and to the various candidates.

Some candidates have greater skill and fortune than others at raising contributions from different sources, or have more resources with which to do so. Some districts are in more expensive media markets and therefore induce a greater need to raise money. Ignoring such variations among candidates induces an artificially high appearance of serial correlation in the contribution series, as does ignoring district-specific effects that persist throughout the campaign. Consequently the effects of lagged and cross-lagged contributions will be overstated. One should therefore take into account district-specific and candidate-specific effects that may persistently increase or decrease the ability of the candidate to raise money.

Other reasons for heterogeneity among candidates may also be important. Historical electoral returns in a district can play a central role both in attracting a particular type of candidate and in setting up expectations for the possibility of a candidate's success. Historical returns are generally time-invariant within a campaign in that they play an important initial role in influencing what types of candidates will contest a district's election. The effect of previous returns may be diminished by unexpectedly good performance in polling or fund-raising by a traditionally disadvantaged party. One should also directly take into account the quality of the candidates as this has been shown to be highly correlated with fund-raising success.

For non-incumbents, previous experience in elected office has been used effectively to measure candidate quality (Jacobson 1990). Ideological positions are another source of candidate heterogeneity. An endorsement by a PAC suggests that the candidate is predisposed to a cluster of issues promoted by the PAC and, more importantly for the current analysis, may signal the attractiveness and viability of a candidate to like minded groups and individuals.

In addition to local effects, there is also reason to believe that contributions are influenced by national considerations. In a previous paper (Wand and Mebane 1999) we analyzed aggregate individual contributions using a model that implies moderating behavior by individual contributors. We found support for the hypothesis that individuals use expectations about the Presidential election outcome when deciding whether to donate money to a House candidate.

Generalized Linear Mixed Models for Contributions

A primary methodological challenge for modeling campaign contributions is to specify a reasonable approximation to the data generating process. The itemized data contain a record for each contribution that a campaign committee reported to the FEC. For contributions from individuals, the recorded value of each contribution has a positive lower bound because small contributions are not itemized: a contribution of less than \$500 from an individual is not reported as an itemized record in the FEC data that we use from 1984. For all contributions the recorded values have an upper bound due to legal limits on how much an individual or PAC may contribute. The legal limits specify restrictions both on the amount someone may contribute to a single campaign and on the total the person may contribute to all campaigns during an election cycle. Moreover, contributions most often occur in one of a few distinct, "round figure" amounts. Such special features of the recorded values make it doubtful that an assumption that contributions are normally distributed is correct. Also doubtful is the modified assumption that the distribution of contributions is normal except for censoring or truncation at the bottom and at the top. Table 1 shows that during the 1984 campaign cycle most contributions from PACs occurred in one of three amounts: \$250, \$500 or \$1,000. Seventy percent of PAC contributions were of one of those sizes. Including other "round figure" amounts—$100, $200, $300, $750, $1,500, $2,000, $2,500 and $5,000—covers 88 percent percent of PAC contributions. The concentration of contributions on a few distinct values is even more stark for contributions from individuals. Two sizes of contributions, $500 and $1,000, account for 93 percent of all donations of $500 or more.

—Table 1 about here—

The results in Table 1 are only suggestive. A better test of the hypothesis that contributions have an additive normal disturbance is to regress the dollar amount of contributions on plausible sets of regressors and then test the distribution of the residuals. We regressed both dollar amounts and logarithms of dollar amounts of contributions on the same variables used in the GLMM specifications that are presented below. For both individual and PAC contributions, tests such as the Kolmogorov-Smirnov test resoundingly reject that the ordinary least squares residuals have normal distributions.

We propose an alternative model of contributions based on a conditional compound Poisson density: conditional on the values of unobserved random effects, contributions are generated by a compound Poisson process.[1] We model the number of contributions received by a candidate during each week and, conditional on the occurrence of a contribution, the size of each contribution in dollars. We assume that given the random effects, the number of contributions is generated by a dynamic Poisson process and the size of each contribution is generated by another process. As we describe in more detail below, for contributions from individuals we specify the size-generating process as a conditional probit model, while for contributions from PACs we assume that the sizes are generated by a conditional Poisson model.

We begin by conceptualizing the number of contributions received by a candidate during a time period as being the result of a count process. This leads to a notable difference in how to interpret the absence of contributions to a candidate during a period of time. A model based on

[1]See Feller (1971, 180–81) for an introductory discussion of compound Poisson processes.

a Gaussian data generating process for the amount of a contribution would suggest that when no contribution is observed, all potential contributors would rather have taken money away from the candidate. In contrast, zero contributions are a natural result of a count process. In a count process, a zero implies no more than that no one wished to make a contribution to the candidate.

Our aim is to model the number of separate actions individuals and PACs took to contribute to a candidate's campaign. Some of the transactions in the itemized FEC data seem to represent decisions individuals made to bundle their contributions to a candidate so they would arrive at the same time. For example, in the record of contributions from individuals we frequently observe spouses on the same day contributing identical amounts to a candidate. Also frequent in the record of individual contributions are apparent clusters in which several individuals who work for the same employer contribute to a candidate on the same day. Corporations bundling contributions and presenting them as a single package is a frequently cited form of gaining influence, without being restricted by contribution limits. An example of an employer-based cluster is shown in Table 2, where seven employees from Mobil Oil contributed a total of $5000 on the same day to Republican candidate Dioguardi (NY20). We treat this cluster as a single contribution of size $5000. In general, we define a cluster to include all contributions that occur within the same week from individuals who either have the same family name and an address in the same city, or have the same employer.[2] Most individuals make contributions that are not related by family name or employer to other contributions during the same week. We therefore treat most contributions as comprising a cluster that contains a single transaction.

—Table 2 about here—

Separate itemized contributions from the same PAC often appear during a short period of time. In particular, when PACs give in-kind contributions they may declare each item separtely. An example of this is also shown in Table 3. During the three day period of May 30 to June

[2]We merge a family and an employer cluster together when there is a transaction that belongs in both.

01, the League of Conservation Voters gave \$149 in in-kind contributions to candidate Evans (WA01, Dem), with eleven separate itemized transactions. Given that our empirical analysis aggregates contributions to weekly counts, we treat these eleven contributions as part of a single cluster. Since we are interested in modeling the number of separate actions to contribute, we count the number of individual and PAC clusters, instead of counting each transaction. Henceforth we use the terms "counting contributions" and "counting clusters" interchangeably.

—Table 3 about here—

Conditional on the occurrence of a contribution, we also model the size of each contribution in dollars. As one would suspect from the previous comments about Table 1, this portion of the density is not continuous. Even with the aggregation across transactions that occurs in some clusters, the contribution data exhibit considerable lumpiness. We treat the contribution sizes as categorical donation levels. Itemized contributions from individuals we separate into two categories: medium (\$500) and large (> \$500).[3] For PAC contributions, approximately 80 percent of the sizes are exact multiples of \$250; we round each size to the nearest \$250. Combining the model of the number of contribution clusters with the model of the value of each cluster, we have a model of campaign contributions that can offer predictions denominated in dollars, while also employing a reasonable set of distributional assumptions in the statistical formulation.

To estimate the conditional compound Poisson model, we adopt the GLMM framework. In the preceding discussion of factors that affect contribution decisions we suggested that there is considerable heterogeneity across candidates and districts. Some variables that affect the mean number of clusters or the size of contributions may not be observed by a researcher. Because we shall assume that the random effects affecting counts are independent of the random effects

[3]Note that although we do not have itemized records for small contributions (< \$500) from individuals, we do know the total amount of the small contributions, from FEC spread file data. Such contributions require special handling that we do not describe or implement in this paper.

affecting cluster sizes, we can estimate the likelihoods for each of these separately. We present the count likelihood and cluster size conditional likelihood in sequence.

In a GLMM specification for a conditional Poisson process, the mean is a log-linear function of the unobserved effects, which are treated as random variables. For a single observed count y_t, the Poisson mean given observed vectors $\mathbf{x}_t$ and $\mathbf{z}_t$, a fixed effect parameter vector β and a vector of unobserved random effects $\mathbf{u}$ is

$$\mu_t = \exp(\mathbf{x}_t'\beta + \mathbf{z}_t'\mathbf{u}) \ .$$

The conditional density for observation y_t is

$$f_{y|u}(y_t \mid \mathbf{u}, \mathbf{x}_t, \mathbf{z}_t, \beta) = e^{-\mu_t}\mu_t^{y_t}/y_t! \ .$$

For a sample of T observations we have the conditional joint density,

$$f_{y|u}(\mathbf{y} \mid \mathbf{u}, \mathbf{X}, \mathbf{Z}, \beta) = \prod_{t=1}^{T} f_{y|u}(y_t \mid \mathbf{u}, \mathbf{x}_t, \mathbf{z}_t, \beta) \ , \tag{1}$$

where $\mathbf{y}$ is the vector of all the observed counts, and $\mathbf{X}$ and $\mathbf{Z}$ are the matrices of variables $\mathbf{x}_t$ and $\mathbf{z}_t$. The specification of (1) says that conditional on all observed and unobserved effects, the counts are independently distributed Poisson random variables.

To complete the GLMM specification one integrates (1) with respect to the distribution assumed for the random effects. When that distribution has a density f_u that is a function of parameters $\mathbf{D}_c$, $\mathbf{u} \sim f_u(\mathbf{u} \mid \mathbf{D}_c)$, we have the likelihood

$$L(\beta, \mathbf{D}_c \mid y, \mathbf{X}, \mathbf{Z}) = \int f_{y|u}(\mathbf{y} \mid \mathbf{u}, \mathbf{X}, \mathbf{Z}, \beta)\, f_u(\mathbf{u} \mid \mathbf{D}_c)\, d\mathbf{u} \tag{2}$$

In general (2) has no closed form solution.

A similar derivation occurs to obtain a GLMM specification for the size of each cluster. Let S_{tk} be the dollar value of the kth cluster's contribution in period t, where $k \in \{1, \ldots, y_t\}$. As noted above, we have chosen to classify each individual contribution cluster using dichotomous categories (medium and large). For PAC clusters, we chose a polytomous categorization, rounding the total value of the contributions in each cluster to the nearest \$250. Let $\tilde{S}_{tk}$

be the transformed amount variable, where for example in the individual contribution case, $\tilde{S}_{tk} \in \{\text{med}, \text{lrg}\}$.

We define a linear predictor that has fixed effects parameters γ and random effects $\mathbf{v}$:

$$\tilde{\eta}_t = \mathbf{x}_t'\gamma + \mathbf{z}_t'\mathbf{v} \ .$$

We assume that $\mathbf{v}$ and $\mathbf{u}$ are independent. The conditional density of $\tilde{S}_{tk}$, given that $y_t > 0$, is denoted by $f_{\tilde{S}|v}(\tilde{S}_{tk} \mid \mathbf{v}, \mathbf{x}_t, \mathbf{z}_t)$ and has a functional form that depends on whether we are estimating cluster sizes of individuals or of PACs. In either case, for a sample of T observations, we have the conditional joint density

$$f_{\tilde{S}|v}(\tilde{\mathbf{S}} \mid \mathbf{v}, \mathbf{y}, \mathbf{X}, \mathbf{Z}, \gamma) = \prod_{\substack{t=1 \\ y_t>0}}^{T} \prod_{k=1}^{y_t} f_{\tilde{S}|v}(\tilde{S}_{tk} \mid \mathbf{v}, \mathbf{x}_t, \mathbf{z}_t, \gamma) \ , \tag{3}$$

where (3) says that the size of each contribution is independently distributed. As we do for counts, we complete the cluster size conditional likelihood by integrating over the random effects:

$$L(\gamma, \mathbf{D}_a \mid \tilde{\mathbf{S}}, \mathbf{y}, \mathbf{X}, \mathbf{Z}) = \int f_{\tilde{S}|v}(\tilde{\mathbf{S}} \mid \mathbf{v}, \mathbf{y}, \mathbf{X}, \mathbf{Z}, \gamma) f_v(\mathbf{v} \mid \mathbf{D}_a) \, d\mathbf{v} \ . \tag{4}$$

The difficulty of solving such integrals over random effects has motivated a number of approaches that involve estimation of models that loosely approximate (2). The generalized estimating equations (GEE) approach (Liang and Zeger 1986, Zeger, Liang and Albert 1988) is based on estimating moments that roughly approximate moments derived from (2). Breslow and Clayton (1993) derive a penalized likelihood method by truncating a Laplace transform of (2). Neither the GEE nor the penalized likelihood methods give consistent estimates for the random effects, and both have exhibited substantial bias in sampling experiments (Breslow and Clayton 1993, Kuk 1995, McCulloch 1997, Jiang 1998).

Estimates from the MCEM algorithm of McCulloch (1997) are maximum likelihood (ML) estimates that are consistent for both the fixed effects and the random effect variance parameters of (2). MCEM estimates have not exhibited substantial bias in sampling experiments. The most important novelty in McCulloch's proposal is use of a Metropolis algorithm to simulate a good sample of draws from the distribution of the random effects. McCulloch's MCEM algorithm is a Monte Carlo Markov Chain algorithm. Booth and Hobert (1999) suggest replacing the Metropolis algorithm with rejection or importance sampling methods, along with other changes to improve the efficiency of the optimizing computations. In all variations of the optimization the simulated sample allows (2) to be evaluated in a straightforward way: by Monte Carlo integration. In the Appendix we outline the MCEM algorithm we use.

For what we refer to as a GLMM Poisson-Normal (GLMM-PN) specification, the density $f_u(\mathbf{u} \mid \mathbf{D}_c)$ is a product of densities for normally distributed variables. When there is a single random effect, it is possible to use quadrature to evaluate the integral in (2).[4] In that case it is possible to estimate the parameters of (2) by ML without introducing the simulation error that the Monte Carlo estimation part of the MCEM algorithm entails. But quadrature evaluation of (2) is generally not feasible for more than one random effect, nor for random effects that are not independently distributed. MCEM is feasible in cases where quadrature-based ML estimation is not.

In the contribution counts data, omitting relevant measures of district and candidate heterogeneity would produce overdispersion in a Poisson regression. In the models we have estimated for this paper, we have included observed variables in order to reduce such potential sources of misspecification. Our inclusion of lagged dollar values, in addition to being of substantive theoretical interest for the question of whether contributions are reactive, is also important to take into account a potential source of contagion across observations. Overdispersion would also result if individuals and PACs contribute in a coordinated fashion in ways that the Poisson model ignores. The appearance of such coordination in the original itemized FEC data is the reason we analyze counts and dollar sizes of clusters of contributions rather than the raw transaction records.

[4] SAS V7 includes the experimental procedure NLMIXED that performs such estimation.

A System of Poisson Processes

We define a model in which, conditional on a set of observed and unobserved variables, the number of contributions in each period is independently Poisson distributed.[5] The model describes separate counts for each of several sources of contributions, sets of candidates of each party who receive the contributions, congressional districts and epochs of time (e.g., each day or week) during a two-year election cycle. The counts are conditionally independent, but correlated in their joint unconditional distribution because of exogenous observed and unobserved variables that commonly affect their conditional Poisson means. The counts are correlated also because the conditional Poisson mean of each count is affected by contributions received at earlier times. The conditional Poisson mean of the count of contributions received from each source by each party's set of candidates in each district is, in general, affected by the dollar values of the contributions received in previous time periods from all sources by all candidates in the same district. In this way the processes for all kinds of contributions in each district form a system of processes that interact dynamically and evolve together over time. We assume that the parameters that characterize the system's dynamics are the same in every district.

A key point about the systems is that their dynamics are not in general stationary. The model specification allows for explosive changes in contributions to one or more of the sets of candidates in a district. Such explosive changes need not occur during the time span of an election cycle, even when the model's parameters describe nonstationary system behavior. In reality, in the FEC data, unbounded explosions of contributions do not occur. But the model can describe cases in which contributions to one party's candidates suddenly increase dramatically while contributions to the other party remain level or even decline. The model can also describe cases in which contributions to all parties' candidates rapidly increase in a reactive frenzy.

Let y_{hijt} be the count of clusters of contributions to a set of House candidates and S_{hijtk} be

[5] Again, when we refer to counts of contributions, we mean the number of clusters.

the dollar value of each cluster, where the indices are used as follows,

$h \in H = \{h_1, \ldots, h_{m_H}\}$ source of contribution (for individual or PAC, $m_H = 2$)

$i \in I = \{i_1, \ldots, i_{m_I}\}$ party (for Democrat or Republican, $m_I = 2$)

$j \in J = \{j_1, \ldots, j_{m_J}\}$ congressional district (to cover all districts, $m_J = 435$)

$t \in T = \{t_1, \ldots, t_{m_T}\}$ epoch within the election cycle (to cover all weeks, $m_T = 104$)

$k \in K = \{1, \ldots, y_{hijt}\}$ distinct clusters

The daily count of contributions from each source to the candidates of each party in each district is independently Poisson distributed, conditional on a set of exogenous variables, previous weekly dollar totals of contributions, fixed effect parameters and random effects. The total amount of money contributed from source h to candidates of party i in district j during time t is

$$S_{hijt} = \begin{cases} 0, & \text{if } y_{hijt} = 0 \\ \sum_{k=1}^{y_{hijt}} S_{hijtk}, & \text{if } y_{hijt} > 0 \,. \end{cases}$$

Using the linear predictor

$$\eta_{hijt} = \mathbf{x}'_{hijt}\beta_h + \left(\sum_{\bar{h} \in H} \sum_{\bar{i} \in I} \sum_{\bar{t}=t-q_{h\bar{h}}}^{t} S_{\bar{h}\bar{i}j\bar{t}} \alpha_{hi\bar{h}\bar{i},t-\bar{t}} \right) + b_{hj} + c_{hij} \,, \qquad (5)$$

the mean of the Poisson density is

$$\mu_{hijt} = \exp(\eta_{hijt}) \,.$$

The variables in the vector $\mathbf{x}_{hijt}$ are assumed to be exogenous. β_h is a vector of fixed effect parameters. The set

$$\mathbf{S}_{hj,<t} = \bigcup_{\bar{h} \in H} \bigcup_{\bar{i} \in I} \bigcup_{\bar{t}=t-q_{h\bar{h}}}^{t} S_{\bar{h}\bar{i}j\bar{t}}$$

contains the lagged contributions pertinent for y_{hijt} from all sources and parties included in the system. The nonnegative integers $q_{h\bar{h}}$ specify the maximum order up to which lagged amounts from source $\bar{h}$ may affect the mean of counts from source h. Each $\alpha_{hi\bar{h}\bar{i},t-\bar{t}}$ is a scalar fixed effect parameter. Let

$$\alpha_{hi} = \bigcup_{\bar{h}\in H} \bigcup_{\bar{i}\in I} \bigcup_{\bar{t}=t-q_{h\bar{h}}}^{t} \alpha_{hi\bar{h}\bar{i},t-\bar{t}}$$

denote the set of such parameters that pertain to y_{hijt}. The parameters in α_{hi} characterize the system's dynamics. In general, the system has nonstationary dynamics if all the values in α_{hi} are positive and stationary dynamics if all the values in α_{hi} are negative. If some values in α_{hi} are positive and some are negative, the stationarity of the overall dynamics depends on the exact combination of the parameter values. α_{hi} is the same for all epochs $t \in T$ and all districts $j \in J$. Let

$$\alpha = \bigcup_{h\in H} \bigcup_{i\in I} \alpha_{hi}$$

denote the set of all such parameters. Any values in the set $\{S_{\bar{h}\bar{i}j\bar{t}} \in \mathbf{S}_{hj,<t} : \bar{t} < t_1\}$ are predetermined initial conditions. Let

$$\mathbf{S}_< = \bigcup_{h\in H} \bigcup_{j\in J} \bigcup_{t\in T} \{S_{\bar{h}\bar{i}j\bar{t}} \in \mathbf{S}_{hj,<t} : \bar{t} < t_1\}$$

denote the set of all such initial conditions. The effects b_{hj} and c_{hij} are random. Defining

$$u_{hijt} = (b_{hj}, c_{hij})' ,$$

we have a conditional Poisson density for each observed count y_{hijt}:

$$f_{y|u}(y_{hijt} \mid u_{hijt}, \mathbf{x}_{hijt}, \mathbf{S}_{hj,<t}, \beta_h, \alpha_{hi}) = \exp\{-\mu_{hijt} + y_{hijt}\eta_{hijt} - \log(y_{hijt}!)\} .$$

Defining the vector of the realizations of all the random effects as

$$\mathbf{u} = (b_{h_1 j_1}, \ldots, b_{h_{m_H} j_{m_J}}, c_{h_1 i_1 j_1}, \ldots, c_{h_{m_H} i_{m_I} j_{m_J}})'$$

implicitly defines $\mathbf{Z}$ of (1) to index the random effects: $\mathbf{z}'_{hijt}\mathbf{u} = b_{hj} + c_{hij}$. The conditional joint density of all the observed counts is

$$f_{y|u}(\mathbf{y} \mid \mathbf{u}, \mathbf{X}, \mathbf{Z}, \mathbf{S}_<, \beta, \alpha) = \prod_{h\in H} \prod_{i\in I} \prod_{j\in J} \prod_{t\in T} f_{y|u}(y_{hijt} \mid u_{hijt}, \mathbf{x}_{hijt}, \mathbf{S}_{hj,<t}, \beta_h, \alpha_{hi}) , \qquad (6)$$

where $\beta = \bigcup_{h \in H} \beta_h$.

We assume all the random effects are normal with zero mean. Each of the effects b_{hj} and c_{hij} is identically and independently distributed with variance, respectively, ϕ_h^2 and σ_h^2. In formal terms,

$$
\begin{aligned}
b_{hj} &\sim f_b(b_{hj} \mid \phi_h^2) = (2\pi\phi_h^2)^{-1/2} \exp\{-b_j^2/(2\phi_h^2)\} \\
c_{hij} &\sim f_c(c_{hij} \mid \sigma_h^2) = (2\pi\sigma_h^2)^{-1/2} \exp\{-c_{hij}^2/(2\sigma_h^2)\} \,.
\end{aligned}
$$

With $\mathbf{D}_c = (\phi_{h_1}^2, \sigma_{h_1}^2, \ldots, \phi_{h_{m_H}}^2, \sigma_{h_{m_H}}^2)'$, the joint density of the realizations of all the random effects is

$$
f_u(\mathbf{u} \mid \mathbf{D}_c) = \left[\prod_{h \in H} \prod_{j \in J} f_b(b_{hj} \mid \phi_h^2)\right] \left[\prod_{h \in H} \prod_{i \in I} \prod_{j \in J} f_c(c_{hij} \mid \sigma_h^2)\right] \,.
$$

The GLMM-PN likelihood takes the form

$$
L(\beta, \alpha, \mathbf{D}_c \mid y, \mathbf{X}, \mathbf{Z}, \mathbf{S}_<) = \int f_{y|u}(\mathbf{y} \mid \mathbf{u}, \mathbf{X}, \mathbf{Z}, \mathbf{S}_<, \beta, \alpha)\, f_u(\mathbf{u} \mid \mathbf{D}_c)\, d\mathbf{u} \,. \tag{7}
$$

The source-and-district-specific effect b_{hj} induces correlation between the counts of contributions to both parties within each district from each source. For example, if a House race is designated by pundits as some kind of bellwether district, both parties may receive increased contributions as partisan or ideological contributors may exert particular effort in order to raise expectations for friendly candidates in other districts. Because the data we use do not include measures of advertising cost within a district, this random effect may also capture the relative difference in the basic need for raising money. For simplicity we refer to b_{hj} as the district-specific effect

The source-and-party-and-district-specific effect c_{hij} induces correlation between the counts of contributions from a source to one party within a district. Parties in some districts may simply be better than other parties at raising money from a source. For simplicity we refer to c_{hij} as the candidates-specific effect.

Cluster Size Models

We use two GLMM specifications for the cluster sizes. Each includes random effects $v_{hijt} = (d_{hj}, e_{hij})'$ that we assume to be identically and independently normal with zero mean and variance, respectively, $\tilde{\phi}_h^2$ and $\tilde{\sigma}_h^2$.

For the dollar sizes of the contributions from PACs we use a GLMM-PN model that has the same form as the model for the number of contributions, except omitting the effects of the lagged dollar amounts. The linear predictor for the Poisson density is

$$\tilde{\eta}_{2ijt} = \mathbf{x}'_{2ijt}\gamma_2 + d_{2j} + e_{2ij} \ . \tag{8}$$

where the variables in the vector $\mathbf{x}_{2ijt}$ are as defined in (5), and γ_2 is a vector of fixed effect parameters. The dependent variable in this model—the "count" for the conditional Poisson likelihood—is defined by rounding the dollar value of each contribution to the nearest \$250 and then dividing by 250.

For the dollar sizes of the contributions from PACs we use a conditional probit model for the choice between the medium and large categories. Again we omit the effects of the lagged dollar amounts. The linear predictor for the probit is

$$\tilde{\eta}_{1ijt} = \mathbf{x}'_{1ijt}\gamma_1 + d_{1j} + e_{1ij} \ . \tag{9}$$

The conditional probit likelihood is $\Pr(\tilde{S}_{1ijt} = \text{lrg} \mid \mathbf{x}_{1ijt}, d_{1j}, e_{1ij}, \gamma_1) = \Phi(\eta_{1ijt})$, where Φ denotes the cumulative normal distribution function.

Data and Detailed Model Specification

The campaign contributions data for individuals and PACs are drawn from the Federal Election Commission (FEC) Itemized Contributions files for 1984. To match the temporal span of key explanatory variables that derive from the American National Election Study (ANES) 1984 Continuous Monitoring Study, described below, we restrict the data to include contributions only during the 42-week period from mid-January 1984 until just before the general election.

We further limit our sample to the twenty open seat races that meet criteria described in the Appendix. The main reason to restrict the analysis to the open seat races is that we think the dynamics of campaign contributions are different in such races than they are when an incumbent is involved. Restricting the analysis also facilitates computation. We aggregate the counts of contributions to the level of parties within districts and to the frequency of weeks. More details on the dataset used for this paper are presented in the Appendix. We define separate series for each party and each source of contribution by district. The first week begins January 16, 1984, and the last begins October 29, 1984. Figures 1 to 4 plot the values of the counts of weekly contributions from the beginning of January, 1984 until the week prior to election day. The variables in **X** are as follows.

—Figures 1 to 4 about here—

I(district) In some of our model specifications we use dummy variables to estimate district-specific fixed effects, omitting the dummy for AL01 (which therefore determines the intercept). In a model that included many more than twenty open seat races, the use of dummy variables would become unattractive for reasons of both statistical consistency and computational stability as number of districts increased. In such a situation, accounting for district heterogeneity by random effects would become much more attractive.

Lagged Amount$_{t-\tau}$ Functions of the lagged total amounts from each of the series are are included in the full system of contribution counts. We argue that the current number of contributions, in part, are influenced by funds recently raised both by your own party and that of the opponent. For example, as discussed above, previously raised funds enables a candidate to undertake a better equipped campaign, and thus raise more money in the future. We allow for the possibility that money from different sources do not have the same impact. PACs may be more sensitive to past individual contributions, while individuals may pay little attention to the source of past fund raising. We also allow a candidate's potential contributors to react to contributions received by the opposing party.

"Own" and "opp" designations in table labels indicate simply whether the lagged contributions $S_{hi'j,t-\tau}$ and the count y_{hijt} relate to the same party, $i' = i$ ("own"), or to the opposing party, $i' \neq i$ ("opp"). This is not meant to suggest any judgment about the preference of the contributors, in terms of defining which party a contributor may perceive as the opposition.

As noted above, small contributions (<\$500) are not itemized in the 1984 individual contributions dataset, so that we do not observe the dates at which they occurred. To ensure we have at least the mean level of dollar totals correct, we multiply the lagged weekly totals of individuals' contributions by the inverse of the ratio of the dollar total of itemized individual contributions to the dollar total of all individual contributions.[6] For example, if a candidate's contributions of \$500 or more constituted one-third of all her contributions, then each weeks lagged values would be multiplied by three. This distributes the value of the small contributions over the entire campaign in proportion to the amounts we do observe during each time period.

To give the point estimates of the α_{hi} parameters a more convenient scale, we divide the lagged amounts by 1000. The unit of measurement for the lagged variables $S_{\bar{h}\bar{i}j\bar{t}}$ is therefore \$1000.

HQ We seek to measure the quality of the candidates, as this is likely to be related both to the beliefs potential contributors have about the candidate's prospects and the candidate's organizational ability to raise money. Quality is a dichotomous variable based solely on whether a candidate has previously held elected office. A non-incumbent who has done so is deemed high quality. The quality dataset was collected by Gary Jacobson; see Jacobson (1990). Because Jacobson's data only indicate the quality of the general election candidates, our classification of some parties as having a high quality open seat candidate is actually an approximation. A candidate who did not survive to the general election may have been of high quality. If a party's general election candidate in a district is designated as high quality in Jacobson's data, then

[6] We use FEC 1983–84 Spread File data (ICPSR 8511) to measure the total dollar value of all contributions.

we designate the entire aggregate contributions Poisson process for that party in the district as having the high quality dummy variable equal to one. The HQ dummy is zero otherwise.

$1/H_t$ is the inverse of the Herfindahl-Hirschmann concentration index, computed based on the primary election vote for each candidate. Using p_i to denote the share of the vote received by candidate i of n_P candidates running in the primary election at time t_P,

$$H_t = \sum_{i=1}^{n_P} p_i^2, \quad t < t_P \ .$$

$1/H_t$ measures the effective number of candidates. For discussion and applications of this index to parties see Taagepera and Shugart (1989, 79). In general H_t varies over the campaign, with the maximum value being achieved during the pre-primary period. During the general election period—after the last primary or run-off vote—we set $H_t = 1$ for each major party. If a run-off election between two candidates was necessary after a non-decisive primary, we set $1/H_t = 2$ up to the time of the run-off.[7] When there is only a single candidate registered with the FEC and that candidate wins the party's nomination by acclamation, $H_t = 1$ for the entire campaign.

This variable taps into two considerations. In terms of the decision to contribute or not, when more candidates are asking for money there are likely to be a greater number of contributions to a party. As the number of candidates grows there may be an incentive to reduce the size of each contribution, either due to hedging by contributors or budget constraints of those who give to multiple candidates.

One Candidate is a dummy variable for a party that had only a single candidate registered with the FEC who won the parties nomination by acclamation. This dummy variable is included since we wish to distinguish between the case where $H_t = 1$ because there is only a single candidate, and $H_t = 1$ because all other candidates have been defeated. A single candidate suggests an uncompetitive race because either the party's chances of winning the general election

[7]We collated the primary and run-off dates and vote totals during 1984 from Scammon and McGillivray (1985).

are very small or the victory of the sole candidate is assured. In the first case contributions are less eagerly made, and in the latter contributions may be less eagerly sought after. Either way, contributions may be lower.

Party We allow the Republican and Democratic parties to have different means on average across the different open seat races by including a dummy variable, Party, which is equal to one if the dependent variable of the series is for contributions to Democrats and zero otherwise.

f(Dem vote 1982) Past Vote measures the level of support for the Democratic House candidate in the 1982 general election. We specify the basic variable as the two-party vote proportion for the Democrat minus 0.5. To allow for opposite, but equal effects, for each party, we multiply the basic variable by 1 for Democratic series and -1 for Republicans.

I($\mathbf{PAC}_t$) We include time-varying dummies for endorsements from six of the most influential PACs in 1984: Americans for Democratic Action (ADA), AFL-CIO, Committee on Political Education (COPE), National Committee for an Effective Congress (NCEC), Americans for Constitutional Action (ACA), Business-Industry PAC (BIPAC), and National Conservative PAC (NCPAC). These PACs are deemed important in large part because their endorsements received so much attention. For example, *Congressional Quarterly Weekly Report* (*CQWR*) published a summary of general election endorsements by these PACs in 1984 (*CQWR* Nov. 17, pp. 2971–76).

The *CQWR* endorsement data have two major limitations: the data deal only with general election candidates and provide no date of the endorsement. Since we know for some PACs that contributions are closely associated with the decision to endorse,[8] we can sometimes use the first expenditure by a PAC on a candidates behalf[9] as a proxy for the decision/announcement date. For BIPAC, NCEC, and NCPAC we use the date of first contribution as the date to turn

[8]We got in touch with officials of most of the PACs to learn how each one handles endorsements.

[9]We treat transaction codes 24E, 24F, 24K, and 24Z as expenditures on a candidates' behalf for this measure.

on the endorsement dummy. In the case where the endorsed candidate was not the general election candidate, and the general election candidate received no funds from the PAC, we turn off the endorsement dummy after the primary (or run-off where appropriate). For COPE and ADA we use the date of first expenditure when applicable but if, as occurs, an endorsement is recorded in *CQWR* and there is no contribution, the dummy is turned on for the entire campaign. For ACA, which made no contributions to the candidates in our data, the *CQWR* endorsements are also used for the entire campaign.

Having information about the timing of endorsements we are able to comment on whether some PAC's endorsements instigated higher contribution levels. In addition, or in the absence of timing information, the endorsement dummy variable accounts for differences in ideology and interests between the candidates in the general election.

$(\theta_{Dt} - \theta_{Rt})$ is the difference between the expected policy locations of the two major parties' presidential candidates. Each policy location is mapped onto the $[0, 1]$ interval, with 0 being most liberal position and 1 being the most conservative. θ_{Dt} is the weighted sum of the locations of each of the Democratic party presidential primary candidates, the weights being each candidate's expected probability of nomination. Both the candidate positions and the candidate's probabilities of winning are derived from data taken from the ANES 1984 Continuous Monitoring Study (Miller and the National Election Studies 1985). The Continuous Monitoring Study performed daily interviews of random cross-sections of individuals from 11 January 1984 through 31 December 1984. θ_{Rt} is the policy location of President Ronald Reagan, the de facto Republican candidate throughout the 1984 primary season; we treat θ_{Rt} as constant throughout the entire year. See Wand and Mebane (1999) for more details about θ_{Dt} and θ_{Rt}.

$\bar{P}_t$ is the average subjective probability of Reagan winning the general election. The expected probability of winning data for Reagan are also derived from the Continuous Monitoring Study. Details regarding the construction of the national-level variables and their analytical motivation within a policy moderating theory of campaign contributions are described in Wand and Mebane (1999). The key predictions from the policy moderating model concern the effects of $(\theta_{Dt} - \theta_{Rt})$ and $\bar{P}_t$. We show in Wand and Mebane (1999) that a policy moderating model predicts that during 1984 the effect of $(\theta_{Dt} - \theta_{Rt})$ on the number of contributions should be positive, likewise the effect of $\bar{P}_t$.

Estimation and Results

We construct a system of conditional compound Poisson processes that includes both individual and PAC contributions. Because the random effects for individual contributions are independent of the random effects for PAC contributions, we may consistently and efficiently estimate the parameters that apply to individual contributions ($h = h_1$) separately from the parameters that apply to PAC contributions ($h = h_2$). The independence between $\mathbf{u}$ and $\mathbf{v}$ and the fact that β and γ do not have any parameters in common allows us separately to estimate the parameters of the models for the number of contributions and for the size of each contribution.

The count models include four weeks of lagged dollar amounts for the party and for the opposition party, treating individual and PAC amounts separately; in terms of (5), $q_{h\bar{h}} = 4$.[10] Within each half of the system we have imposed symmetry conditions on the parameters of the dynamic relations: for i_1 = Democrat and i_2 = Republican, we impose the equalities $\alpha_{hi_1\bar{h}i_2,t-\bar{t}} = \alpha_{hi_2\bar{h}i_1,t-\bar{t}}$ and $\alpha_{hi_1\bar{h}i_1,t-\bar{t}} = \alpha_{hi_2\bar{h}i_2,t-\bar{t}}$. The first equality assumes that contributions to both parties' candidates react to previous fundraising by the opposing parties' candidates with exactly the same effect. The second equality assumes that contributions to both parties' candidates are influenced by their own previous fundraising with exactly the same effect.

We begin by considering what happens if we use the counts of raw recorded transactions for

[10]We tested the models with up to eight lagged weeks and there is a precipitous decline in significant effects after week four. For all lagged variables we use data from the weeks before the time period analyzed in this paper as initial conditions.

the individual contribution series instead of the cluster counts. Table 5 presents results from a plain Poisson model and from two GLMM-PN models that have one random effect estimated using quadrature. The random effect is the candidates-specific effect denoted c_{hij} in (5). The effect takes 40 distinct values in the data for each series, one for each party in each district. In the models of Table 5, the district-specific effects b_{hj} of (5) are estimated as fixed effects, by including a dummy variable for each district.

—Table 5 about here—

We first look to the two models that use raw counts and consider the difference between Poisson regression and accounting for heterogeneity with GLMM-PN. We focus primarily on changes in inference and defer more in-depth substantive interpretations until we consider the final model. The two models have noticeably different estimates for the fixed effects of many of the variables that do not vary over time. The standard errors (SEs) of the GLMM-PN model estimates for those effects are larger than the SEs for those same effects in the Poisson model. Several of the GLMM-PN model SEs are nearly three times larger than the Poisson model SEs. The larger SEs of the GLMM-PN estimates are a consequence of the fact that the c_{hij} random effect in the GLMM-PN model varies over districts but does not vary over time. The GLMM-PN model averages over all possible values those effects might take; viz. the integral in (2). The uncertainty about those unobserved values causes the effects of the observed time-invariant variables to be estimated with less precision. The greater the variation in the random effects, the less precise are the inferences one may make about fixed effects associated with observed variables that vary over the same indices as the random effects. If the random effects actually exist, then the estimates from the Poisson model are overstating the precision with which the effects of time-invariant variables may be estimated using the observed data.

Conversely, we can expect to have greater confidence across models in inferences based on the dynamic and time-varying parameters. We note that the estimated effects of the lagged dollar amount variables are generally of smaller magnitude in the GLMM-PN model. The SEs are the same and inferences about the effects would be substantively the same in both models, with few exceptions.

Contrasting the GLMM-PN models of raw counts and cluster counts as dependent variables, there are a few changes in inference among the non-dynamic variables. The key difference is the twenty percent drop in σ^2, the variance of the candidates-specific random effect. Using cluster counts avoids treating overdispersion due to contagion as a component of heterogeneity among the sets of candidates from each party in each district. But even with all the fixed effects that are included in the GLMM-PN model and the use of cluster counts, the variance estimated for the random effect in the third model in Table 5 is significant.

Treating both b_{hj} and c_{hij} as random, we estimate the model using the MCEM algorithm. Results are presented in Table 6. For contributions from individuals the random effects vary more over districts than they do over sets of candidates. For contributions from PACs the opposite is true, with the candidates-specific variance being much larger than the district-specific variance.

The general story about the dynamics is that contributions from individuals react to contributions to the opposing party but contributions from PACs do not. Both series react to previous contributions to candidates of the same party. The greater the dollar amount of contributions a party's candidates received from individuals in each of the previous four weeks, the greater the mean number of contributions the party's candidates receive from both individuals and PACs during the current week. The mean of contributions to a party's candidates from PACs also uniformly increases with greater numbers of previous contributions to those candidates from PACs. But the mean of contributions to a party's candidates from individuals exhibits a complex oscillation in response to greater numbers of previous contributions to the candidates from PACs: the mean of contributions from individuals increases in response to more contributions made by PACs one week and three weeks previously, but decreases in response to PAC contributions from two weeks previously.[11] Contributions to a party's candidates from

[11]The pattern resembles what would be considered a damped sinusoid in a scalar linear Gaussian time series.

individuals react strongly to contributions that candidates of the opposing party receive from both individuals and PACs, although only the second lagged week is significant and that effect is positive. Contributions from PACs do not respond significantly to contributions to the opposing party's candidates at all.

The electoral history of the district strongly affects both individual and PAC contributions. The positive coefficients for f(Dem vote 1982) mean that, other things equal, a larger vote share for a party's candidate in the district in 1982 is followed by the party's candidates in that district being likely to receive more contributions from individuals in each week during 1984, while the other party's candidates in the district are likely to receive fewer contributions. The positive effect of the Party variable suggests that, other things equal, Democrats tend to attract a larger number of contributions from both individuals and PACs. A greater effective number of candidates ($1/H_t$) significantly increases the number of contributions from both PACs and individuals.

Endorsements affect contributions from individuals and contributions from PACs differently. The conservative groups, ACA, BIPAC, and NCPAC, have the greatest success in spurring PAC contributions. But of this group only BIPAC has a significant effect on individual contributions. COPE is the only liberal PAC to have a significant effect on either PACs or individuals. A COPE endorsement tends to increase the number of contributions from both sources.

The national time-series variables that pertain to the policy moderating model have large and significant effects in both series. The policy difference variable ($\theta_{Dt} - \theta_{Rt}$) has a larger effect on the mean number of PAC contributions that it does on the mean number of contributions from individuals. The effect of subjective expectations that Reagan would win the general election ($\bar{P}_t$) affects the mean number of contributions from both types of sources equally. As the policy moderating model predicts, all these effects are positive.

Table 7 reports estimates for the cluster size models. The individual cluster size model includes the smallest number of significant parameters. The variance of the candidates-specific random effect proved insignificant to the point of having a numerically zero point estimate.

The model reported in Table 7 includes only the district-specific random effect. Other parameters that have blank entries in Table 7 had insignificant estimates when included in the model. They are also omitted from the model reported in Table 7. The effect of $1/H_t$ is the only substantively interesting parameter to achieve unambiguous significance: as the effective number of candidates increases, the value of each contribution decreases. This may suggest that individual contributors reduce the size of their contributions until there is a single post-primary candidate. Perhaps they are hedging. Among endorsements, the effects of ACA, COPE and NCPAC endorsements are borderline significant. Other things equal, the ACA and NCPAC endorsements diminish the mean size of a contribution while a COPE endorsement increases it. Recall that neither ACA nor NCPAC had a significant effect on the number of individual contributions, so their overall effect on the coffers of an endorsed candidate appears to be negative.

The PAC cluster size GLMM-PN model with two random effects has a wider array of significant effects. Other things equal, Republicans tend to receive larger contributions. Other things equal, each contribution to a party's candidates in a district tends to be slightly larger when the candidate that becomes the nominee for the general election has previously held elective office (i.e., $HQ = 1$). Other things equal, a larger vote share for a party's candidate in the preceding general election tends to reduce the size of each contribution to the party's candidates in the current race, while the other party's candidates in the district are likely to receive larger contributions. Other things equal, a larger effective number of candidates tends to reduce the size of each contribution. Of the endorsements only endorsements by ACA or BIPAC have significant effects. Other things equal, both endorsements tend to make each contribution smaller. Recall that both ACA and BIPAC had significant positive effects on the number of PAC contributions, so the net effect in terms of total dollars is unclear. The effects of expectations about the presidential election outcome are also significant. The size of a PAC contribution increases both as the average subjective probability of Reagan winning the general election increases and as the difference in the expected policy location of the two major parties' presidential candidates increases. Both of the random effect variances in the model are significantly positive, with the candidates-specific variation being greater than the variation over districts.

Discussion and Conclusions

Using the general and flexible framework of a conditional compound Poisson process to model campaign contributions data, we are able to account for unobserved heterogeneity and dynamic system effects. The estimates of the random effect variances show that even when a number of observable variables that vary over candidates and districts and that significantly affect contributions are included as exogenous regressors, there are still significant unobserved candidates-specific and district-specific effects. Failing to account for such unobserved effects would lead to overestimates of the magnitude of the effects that observed time-invariant variables have on contributions. The dynamics of lagged and cross-lagged contributions are highly significant. Estimates of the dynamic effects do not change materially if the heterogeneity is handled differently, or even if it is ignored. The interpretation of the dynamic coefficients, particularly those that suggest oscillation in the interactions between PAC and individual contributions, will be clearer once simulations of the system's behavior have been completed. Simulations should clarify how close the contributions process is to having explosively nonstationary dynamics.

A limitation of the current analysis is our inability to distinguish between changes in candidate effort and changes in contributor receptivity when an opponent successfully raises contributions. For example, with an increase in the number of contributions to one party and a positive coefficient on an opponent lagged count variable, there are at least two possible interpretations for an increase in the current contribution to the opposing party. First, the increase in contributions leads the other party's candidates to increase their effort to solicit funds. Alternatively, potential contributors may be more receptive to existing solicitations when the opposition is successfully mobilizing funds. With the limited information available to individual contributors, candidates would nevertheless have to make the effort to inform individuals and solicit contributions under the second scenario. We can only estimate the net effect of these changes, with each scenario being partially true.

Appendix: MCEM

We outline the MCEM algorithm we use to estimate our models. Step 2 is taken directly from the MCEM algorithm of McCulloch (1997). Steps 1, 3 and 4 are our inventions.

1. To estimate models with two random effects we choose starting values for β by estimating a GLMM-PN model using the SAS procedure NLMIXED.

2. The MCEM algorithm uses a Metropolis algorithm to get random draws from the conditional distribution of $\mathbf{u} \mid \mathbf{y}$. Iteration m of the MCEM algorithm begins with the generation of N replications of the random effects, $\tilde{\mathbf{u}}^{(1)}, \ldots, \tilde{\mathbf{u}}^{(N)}$, from the candidate distribution $f_u(\mathbf{u} \mid \mathbf{D}^{(m-1)})$. The Metropolis step compares each element of the newly generated replications to the corresponding element of the set of simulated values from the previous iteration, namely, $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(N)}$. The Metropolis algorithm probabilistically replaces the previous element with the new element, using probabilities determined by an acceptance function, A_k. Let $\tilde{u}_k^{(j)}$ denote the kth element of the jth replication $\tilde{\mathbf{u}}^{(j)}$ in the newly generated values. Define $\mathbf{u}^{(j)*} = \left(u_1^{(j)}, u_2^{(j)}, \ldots u_{k-1}^{(j)}, \tilde{u}_k^{(j)}, \ldots, u_q^{(j)}\right)$. One accepts $\tilde{u}_k^{(j)}$ (and replaces u_k within the jth replication $\mathbf{u}^{(j)}$) with probability,
$$A_k(\mathbf{u}^{(j)}, \mathbf{u}^{(j)*}) = \min\left\{1, \frac{f_{y|u}(\mathbf{y} \mid \mathbf{u}^{(j)*}, \mathbf{X}, \mathbf{Z}, \beta)}{f_{y|u}(\mathbf{y} \mid \mathbf{u}^{(j)}, \mathbf{X}, \mathbf{Z}, \beta)}\right\} .$$
A special virtue of A_k is that it depends only on the conditional distribution of $\mathbf{y} \mid \mathbf{u}^{(j)}$. Repeating this acceptance/rejection step for each of the q elements of each $\mathbf{u}^{(j)}$, $j = 1, \ldots, N$, produces the set of simulations of the random effects to be used in the maximization step of iteration m.

(a) Choose $\beta^{(m)}$ to maximize the empirical expectation of the conditional log-likelihood of $\mathbf{y}$:

$$\frac{1}{N}\sum_{j=1}^{N}\log f_{y|u}(\mathbf{y}\mid\mathbf{u}^{(j)},\mathbf{X},\mathbf{Z},\beta)$$

(b) Choose $\mathbf{D}^{(m)}$ to maximize the empirical expectation of the log-likelihood of $\mathbf{u}$:

$$\frac{1}{N}\sum_{j=1}^{N}\log f_u(\mathbf{u}^{(j)}\mid\mathbf{D})$$

(c) Set $m = m + 1$.

3. Iterate until $\beta^{(m)}$ and $\mathbf{D}^{(m)}$ converge. To the extent allowed by computing resources, we rapidly increase N over iterations. We do this in a rigid way, not based on formal estimates of Monte Carlo error as in the algorithms of Booth and Hobert (1999). When growing N we compute the additional random effect replications using a t-distribution with moments computed from the previous set of replications; this technique is inspired by Booth and Hobert (1999)'s importance sampling method. We continue iterations until the $\mathbf{D}^{(m)}$ values suggest that the Monte Carlo Markov Chain has converged to its stationary distribution. At that point, in our data, the $\beta^{(m)}$ values usually have converged to two significant figures. We treat the converged values as the MLEs $\hat{\beta}$ and $\hat{\mathbf{D}}$.

4. Using the MLEs, the variance of the estimates may be estimated using a standard result for random variables w and v:

$$\mathrm{var}(w) = \mathrm{E}_v[\mathrm{var}(w\mid v)] + \mathrm{var}_v[\mathrm{E}(w\mid v)]\ .$$

Setting $w = \hat{\beta}$ and $v = \mathbf{u}$, the asymptotic covariance matrix of $\hat{\beta}$ may be estimated by

$$\mathrm{var}(\hat{\beta}) = \frac{1}{N}\sum_{j=1}^{N}\mathrm{var}\left(\hat{\beta}^{(j)}\right) + \frac{1}{N}\sum_{j=1}^{N}\left(\hat{\beta}^{(j)}-\hat{\beta}\right)\left(\hat{\beta}^{(j)}-\hat{\beta}\right)'$$

where $\hat{\beta}^{(j)}$ is the conditional ML estimate for the jth replication $\mathbf{u}^{(j)}$. The asymptotic variance of the estimated variance component for an identically and independently distributed random effect with q clusters (i.e., an effect that takes q distinct values in the data) may be estimated similarly. Using σ^2 to denote the variance component,

$$\mathrm{var}(\hat{\sigma}^2) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{q}\left[\frac{1}{q}\sum_{k=1}^{q}\left(u_k^{(j)}\right)^4 - \left(\hat{\sigma}^{2(j)}\right)^2\right] + \frac{1}{N}\sum_{j=1}^{N}\left(\hat{\sigma}^{2(j)} - \hat{\sigma}^2\right)^2$$

where

$$\hat{\sigma}^{2(j)} = \frac{1}{q}\sum_{k=1}^{q}\left(u_k^{(j)}\right)^2 .$$

Appendix. Data definitions and selection criteria

We obtained the raw itemized contributions data files from the FEC by anonymous ftp from `ftp.fec.gov:/FEC/`. In combination with the Committee Master files and Candidate Master files, we extracted for each election cycle all new contributions for candidates competing in the upcoming election. Contributions are defined as those records with transaction code 15 for individuals and either code 24K or 24Z for PACs. Itemized individual contributions are included in the FEC data for the 1984 election cycle only for amounts of \$500 or more. Since we are interested in the act of contribution, we exclude refunds which are indicated by new negative contributions.

Our selection includes all contributions to a candidate in the current cycle who has raised or spent \$5,000 toward the current election and any individual who appears on a ballot or who has registered with the FEC but has not yet raised \$5,000.

We selected all districts which were reported to be open seat races early in 1984 by Congressional Quarterly Weekly Report (24 Feb, p.344), and were in fact open seat races at the time of election. Therefore we exclude those open seat races where the incumbent dies during 1984 or lost in the primary. The included districts are: AL01, AR02, CO03, IA05, IL13, IL14, IL22, KS03, MA05, NC09, NH01, NY20, NY30, TN06, TX06, TX19, TX22, UT02, VA07, WA01. We include all candidates who were on the primary ballot, and those candidates who had a valid campaign committee registered with the FEC but who were not the incumbent. We aggregate the counts of contributions to the level of parties within districts and to the frequency of weeks. The time span of our analysis starts during the week of January 16th and ends with the week of October 29th, 1984. Lagged variables for initial weeks of the series (used as the initial conditions) are drawn from weeks prior to the beginning of this time span.

References

Biersack, Robert, Paul S. Herrnson and Clyde Wilcox. 1993. "Seeds For Success: Early Money in Congressional Elections." *Legislative Studies Quarterly* 18:535–51.

Biersack, Robert, Paul S. Herrnson and Clyde Wilcox, eds. 1994. *Risky Business? PAC Decisionmaking in Congression Elections.* Armonk, NY: M. E. Sharpe.

Booth, JamesG. and JamesP. Hobert. 1999. "Maximizing Generalized Linear Mixed Model Likelihoods with an Automated Monte Carlo EM Algorithm." *Journal of the Royal Statistical Society,* B 61:265–85.

Box-Steffensmeier, Janet M. 1996. "A Dynamic Analysis of the Role of War Chests in Campaign Strategy." *American Journal of Political Science* 40:352–71.

Breslow, N. E. and D. G. Clayton. 1993. "Approximate Inference in Generalized Linear Mixed Models." *Journal of the American Statistical Association* 88:9–25.

Feller, William 1971. *An Introduction to Probability Theory and Its Applications*, Vol. II. 2d ed. New York: Wiley.

Himmelberg, Charles P. and Gregory J. Wawro. 1998. "A Dynamic Panel Analysis of Campaign Contributions in Elections for the U.S. House of Representatives." Paper presented at the 1998 Political Methodology Summer Conference.

Jacobson, Gary C. 1990. *The Electoral Origins of Divided Government: Competition in the U.S. House Elections 1946–1988.* Boulder, Co.: Westview Press.

Jiang, Jiming. 1998. "Consistent Estimators in Generalized Linear Mixed Models." *Journal of the American Statistical Association* 93:720–29.

Krasno, Jonathan S., Donald P. Green and Jonathan A. Cowden. 1994. "The Dynamics of Campaign Fundraising in House Elections." *Journal of Politics* 56(2):459–74.

Kuk, Anthony Y. 1995. "Asymptotically Unbiased Estimation in Generalized Linear Models with Random Effects." *Journal of the Royal Statistical Society,* B 57(2):395–407.

Liang, Kung-Yee and Scott L. Zeger. 1986. "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika* 73:13–22.

McCulloch, Charles E. 1997. "Maximum Likelihood Algorithms for Generalized Linear Mixed Models." *Journal of the American Statistical Association* 92:162–170.

Miller, Warren E. and the National Election Study. 1985. "American National Election Study, 1984: Continuous Monitoring Survey." [computer file] Ann Arbor, MI: Center for Political Studies, University of Michigan [original producer]. ICPSR Study 8298, First ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor].

Sabato, Larry J. 1984. *PAC Power: Inside the World of Political Action Committees.* New York: W. W. Norton.

Scammon, Richard M. and Alice V. McGillivray, eds. 1985. *America Votes 16: A Handbook of Contemporary American Election Statistics, 1984.* Washington, D.C.: Congressional Quarterly.

Taagepera, Rein and Matthew Soberg Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems.* New Haven: Yale University Press.

Wand, Jonathan and Walter R. Mebane, Jr. 1999. "Learning in Campaigns: A Policy Moderating Model of Individual Contributions to House Candidates." Paper presented at the 1999 Annual Meeting of the Midwest Political Science Association, April 15-17, Palmer House Hilton, Chicago, IL, Political Methodology Section.

Zeger, Scott L., Kung-Yee Liang and Paul S. Albert. 1988. "Models for Longitudinal Data: A Generalized Estimating Equation Approach." *Biometrics* 44:1049–1060.

Table 1: Most Frequent Itemized Contribution Levels (in Dollars) to House Candidates by Source in the 1984 Election Cycle

PAC

Contribution level	N	Percentage
100	3876	3.0
200	4799	3.7
250	40871	31.5
300	5116	3.9
500	34256	26.4
750	1444	1.1
1000	15764	12.2
1500	1575	1.2
2000	3142	2.4
2500	1759	1.4

Individual

Contribution level	N	Percentage
500	27518	53.5
1000	20526	39.9

Table 2: Example of A Cluster of Contributions, Recipient is Joseph J Dioguardi, (NY20, Rep) on January 1, 1984

Contributor	Residence	Employer	Amount ($)
Murray, Allen	Syosset, NY	Mobil Oil	1000
Tucker, Richard	Westport, CT	Mobil Oil	500
Warner, Rawleigh	New Canaan, CT	Mobil Oil	1000
Schmertz, Herbert	New York, NY	Mobil Oil	1000
Wolfe, Paul J	New Canaan, CT	Mobil Oil	500
Tavoulareas, W P	Sands Pt, NY	Mobil Oil	500
Riordan, James	Greenwich, CT	Mobil Oil	500

Table 3: In-kind Contributions from League Of Conservation Voters to Brock Evans (WA01, Dem)

Date	Amount ($)
May 30, 1984	7
May 30, 1984	7
May 30, 1984	17
May 31, 1984	3
May 31, 1984	15
May 31, 1984	16
May 31, 1984	41
June 01, 1984	2
June 01, 1984	3
June 01, 1984	12
June 01, 1984	26

Table 4: Contribution Size (in Dollars) from PACs to House Candidates Contesting Open Seats in the 1984 Election Cycle, aggregated to weekly totals and rounded to nearest $250

Contribution level	N	Percentage
0	450	7.1
250	1782	28.0
500	1891	29.7
750	143	2.2
1000	1184	18.6
1250	20	0.3
1500	121	1.9
1750	6	0.1
2000	245	3.8
2250	5	0.1
2500	184	2.9
2750	3	0.0
3000	76	1.2
3250	3	0.0
3500	6	0.1
3750	1	0.0
4000	30	0.5
4500	7	0.1
4750	2	0.0
5000	199	3.1
5750	1	0.0
6000	2	0.0
10000	5	0.1

Table 5: Poisson and Poisson-Normal Models of Itemized Individual Contributions using Raw and Cluster Counts as Dependent Variables (continued on the next page)

	Poisson Raw Count		GLMM-PN Quadrature Raw Count		GLMM-PN Quadrature Cluster Count	
	Est.	SE	Est.	SE	Est.	SE
Ind. amt_{t-1}	0.0078	0.0012	0.0064	0.0012	0.0085	0.0013
Ind. amt_{t-2}	0.0082	0.0012	0.0067	0.0013	0.0086	0.0014
Ind. amt_{t-3}	0.0085	0.0012	0.0070	0.0012	0.0072	0.0014
Ind. amt_{t-4}	0.0079	0.0012	0.0066	0.0012	0.0044	0.0014
Ind. opp. amt_{t-1}	0.00014	0.0016	0.0014	0.0016	−0.00045	0.0018
Ind. opp. amt_{t-2}	0.0091	0.0015	0.0100	0.0015	0.0084	0.0017
Ind. opp. amt_{t-3}	−0.0062	0.0018	−0.0041	0.0018	−0.0037	0.0019
Ind. opp. amt_{t-4}	−0.0036	0.0017	−0.0026	0.0017	−0.0020	0.0019
PAC amt_{t-1}	0.0077	0.0012	0.0082	0.0012	0.0078	0.0013
PAC amt_{t-2}	−0.0065	0.0015	−0.0064	0.0015	−0.0055	0.0016
PAC amt_{t-3}	0.0066	0.0014	0.0063	0.0014	0.0060	0.0015
PAC amt_{t-4}	−0.0039	0.0014	−0.0034	0.0014	−0.0022	0.0015
PAC opp. amt_{t-1}	0.00045	0.0013	−0.00029	0.0013	−0.0013	0.0014
PAC opp. amt_{t-2}	0.0053	0.0014	0.0045	0.0014	0.0039	0.0015
PAC opp. amt_{t-3}	−0.0010	0.0015	−0.0017	0.0015	−0.0013	0.0016
PAC opp. amt_{t-4}	0.0025	0.0013	0.0017	0.0013	0.0021	0.0014
Party	0.47	0.074	0.67	0.22	0.66	0.20
HQ	0.13	0.072	0.25	0.24	0.38	0.22
f(Dem vote 1982)	0.45	0.094	0.66	0.30	0.46	0.28
One Candidate	−0.78	0.11	−0.83	0.35	−0.74	0.32
$1/H_t$	0.32	0.019	0.39	0.022	0.38	0.024
I(ACA)	−0.0061	0.070	−0.0081	0.24	−0.088	0.22
I(BIPAC_t)	0.20	0.059	0.20	0.068	0.25	0.073
I(NCEC_t)	0.29	0.076	0.19	0.10	0.20	0.11
I(COPE_t)	−0.12	0.058	0.30	0.079	0.21	0.086
I(NCPAC_t)	−0.32	0.056	−0.18	0.062	−0.14	0.065
I(ADA_t)	−0.055	0.072	−0.064	0.23	0.054	0.21
$\theta_{Dt} - \theta_{Rt}$	8.06	2.56	8.15	2.56	6.96	2.72
$\bar{P}_t$	5.98	1.05	5.81	1.05	5.02	1.12
σ^2			0.098	0.032	0.078	0.028

Table 5: (continued) Poisson and Poisson-Normal Models of Itemized Individual Contributions using Raw and Cluster Counts as Dependent Variables

	Poisson Raw Count		GLMM-PN Quadrature Raw Count		GLMM-PN Quadrature Cluster Count	
	Est.	SE	Est.	SE	Est.	SE
Intercept	−1.85	0.91	−2.11	0.94	−2.07	1.00
I(AR02)	0.20	0.091	0.10	0.38	0.053	0.35
I(CO03)	−1.02	0.11	−1.36	0.38	−1.33	0.34
I(IA05)	−2.10	0.13	−1.98	0.34	−1.86	0.32
I(IL13)	−1.76	0.13	−2.35	0.37	−2.20	0.34
I(IL14)	−1.84	0.13	−2.15	0.37	−2.03	0.34
I(IL22)	−1.95	0.14	−1.92	0.37	−1.86	0.34
I(KS03)	−0.91	0.081	−0.98	0.35	−0.92	0.32
I(MA05)	−0.0059	0.082	0.039	0.36	0.053	0.33
I(NC09)	−0.43	0.078	−0.41	0.35	−0.40	0.32
I(NH01)	−1.85	0.11	−1.83	0.38	−1.76	0.35
I(NY20)	−0.21	0.083	−0.050	0.36	0.022	0.33
I(NY30)	−0.93	0.19	−1.00	0.52	−1.17	0.48
I(TN06)	−0.74	0.092	−0.54	0.36	−0.50	0.33
I(TX06)	−0.40	0.078	−0.46	0.33	−0.36	0.30
I(TX19)	−0.37	0.082	−0.21	0.35	−0.0066	0.32
I(TX22)	−0.85	0.087	−0.97	0.33	−0.81	0.31
I(UT02)	−1.23	0.11	−1.10	0.40	−1.10	0.36
I(VA07)	−0.67	0.097	−0.51	0.37	−0.48	0.33
I(WA01)	−1.43	0.10	−1.40	0.37	−1.44	0.34
I(Mar)	0.14	0.081	0.14	0.081	0.17	0.086
I(Apr)	0.12	0.071	0.11	0.071	0.11	0.075
I(May)	0.093	0.078	0.071	0.079	0.082	0.084
I(Jun)	0.39	0.077	0.40	0.078	0.40	0.083
I(Jul)	0.36	0.079	0.36	0.080	0.33	0.085
I(Aug)	0.48	0.11	0.47	0.11	0.46	0.12
I(Sep)	0.43	0.12	0.44	0.12	0.45	0.13
I(Oct)	0.37	0.13	0.42	0.13	0.39	0.14
−2 Loglikelihood	8381.3		8268.1		7409.9	

Note: $m_H = 1$, $m_I = 2$, $m_J = 20$, and $m_T = 42$, therefore total number of observations is 1680.

Table 6: GLMM-PN with District and Candidate Random Effects: System of Itemized PAC and Individual Contribution Counts

	Individual Series:		PAC Series:	
	Est.	SE	Est.	SE
Intercept	−1.9	1.0	−1.2	1.032
Ind. amt_{t-1}	0.0085	0.0018	0.0051	0.0016
Ind. amt_{t-2}	0.0086	0.0017	0.0057	0.0016
Ind. amt_{t-3}	0.0072	0.0017	0.0072	0.0016
Ind. amt_{t-4}	0.0043	0.0017	0.0050	0.0016
Ind. opp. amt_{t-1}	−0.0003	0.0020	0.0026	0.0017
Ind. opp. amt_{t-2}	0.0084	0.0019	−0.00048	0.0018
Ind. opp. amt_{t-3}	−0.0034	0.0021	0.0014	0.0019
Ind. opp. amt_{t-4}	−0.0019	0.0020	−0.0021	0.0020
PAC amt_{t-1}	0.0078	0.0014	0.0076	0.0010
PAC amt_{t-2}	−0.0056	0.0016	0.0027	0.00093
PAC amt_{t-3}	0.0059	0.0015	0.0020	0.00088
PAC amt_{t-4}	−0.0022	0.0015	−0.00089	0.00085
PAC opp. amt_{t-1}	−0.0014	0.0015	−0.0011	0.0010
PAC opp. amt_{t-2}	0.0037	0.0016	−0.00040	0.0011
PAC opp. amt_{t-3}	−0.0014	0.0016	0.0014	0.0010
PAC opp. amt_{t-4}	0.0020	0.0016	0.00035	0.00092
Party	0.63	0.19	0.32	0.15
HQ	0.040	0.16	0.00015	0.12
f(Dem vote 1982)	0.80	0.20	0.71	0.16
One Candidate	−0.73	0.22	−0.032	0.17
$1/H_t$	0.38	0.045	0.13	0.032
I(ACA)	−0.044	0.17	0.32	0.13
I(BIPAC_t)	0.26	0.15	0.51	0.097
I(NCEC_t)	0.13	0.17	0.057	0.12
I(COPE_t)	0.26	0.14	0.21	0.10
I(NCPAC_t)	−0.13	0.15	0.35	0.12
I(ADA_t)	−0.0095	0.20	0.15	0.16
$\theta_{Dt} - \theta_{Rt}$	8.4	2.8	13.9	3.1
$\bar{P}_t$	5.1	1.1	5.9	1.0
I(Mar)	0.14	0.087	0.33	0.11
I(Apr)	0.089	0.081	0.38	0.10
I(May)	0.060	0.093	0.31	0.11
I(Jun)	0.39	0.099	0.55	0.11
I(Jul)	0.30	0.12	0.74	0.11
I(Aug)	0.41	0.13	0.91	0.14
I(Sep)	0.40	0.15	1.26	0.15
I(Oct)	0.34	0.16	1.2	0.16
σ^2	0.21	0.064	0.32	0.12
ϕ^2	0.64	0.23	0.064	0.040
−2 Loglikelihood	7315.4		6774.2	

Note: $m_H = 2$, $m_I = 2$, $m_J = 20$, and $m_T = 42$, therefore total number of observations is 3360.

Table 7: GLMMs for Cluster Sizes

	Individuals (probit-normal):		PACs (GLMM-PN):	
	Est.	SE	Est.	SE
Intercept	0.18	0.067	1.35	0.58
Party			−0.18	0.050
HQ			0.091	0.032
f(Dem vote 1982)			−0.15	0.054
One Candidate	0.034	0.11	−0.021	0.046
$1/H_t$	−0.057	0.018	−0.030	0.013
I(ACA)	−0.10	0.055	−0.096	0.038
I(BIPAC_t)	0.020	0.064	−0.043	0.028
I(NCEC_t)	−0.13	0.093	0.030	0.038
I(COPE_t)	0.14	0.073	−0.017	0.031
I(NCPAC_t)	−0.097	0.059	−0.037	0.034
I(ADA_t)	−0.019	0.078	−0.011	0.046
$\theta_{Dt} - \theta_{Rt}$			4.5	1.8
$\bar{P}_t$			2.0	0.52
I(Mar)			−0.15	0.060
I(Apr)			−0.37	0.053
I(May)			−0.13	0.055
I(Jun)			−0.31	0.056
I(Jul)			−0.25	0.060
I(Aug)			−0.013	0.075
I(Sep)			−0.28	0.081
I(Oct)			−0.38	0.082
$\tilde{\sigma}^2$			0.0092	0.0028
$\tilde{\phi}^2$	0.016	0.0086	0.0061	0.0025
−2 Loglikelihood	6845.4		36378.3	
N	4991		6195	

Note: Parameters omitted from the individual cluster size model had insignificant estimates when included. They were not estimated in the model shown here.

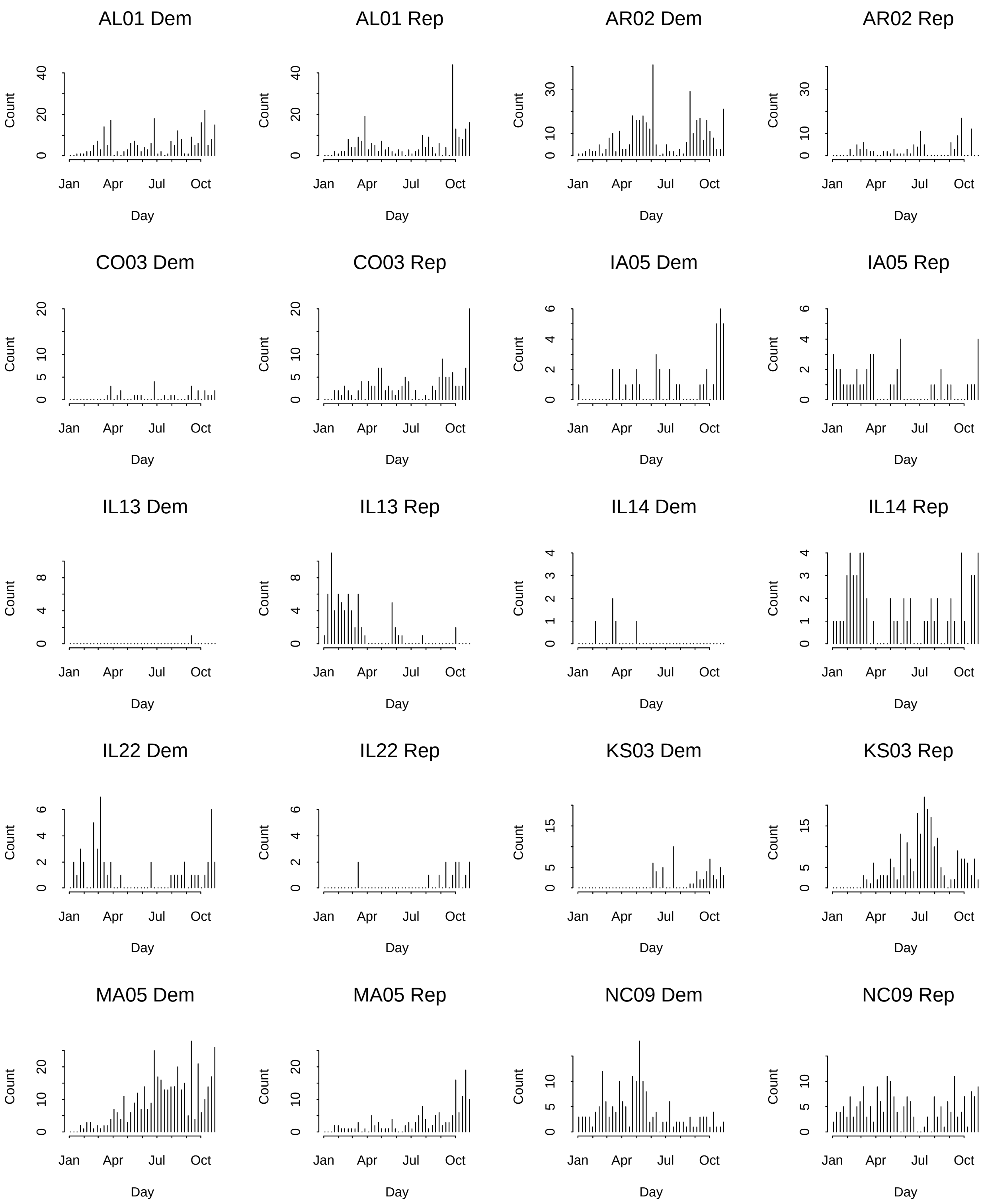

Figure 1: Weekly Counts of Clusters of Contributions to House Candidates: Individuals

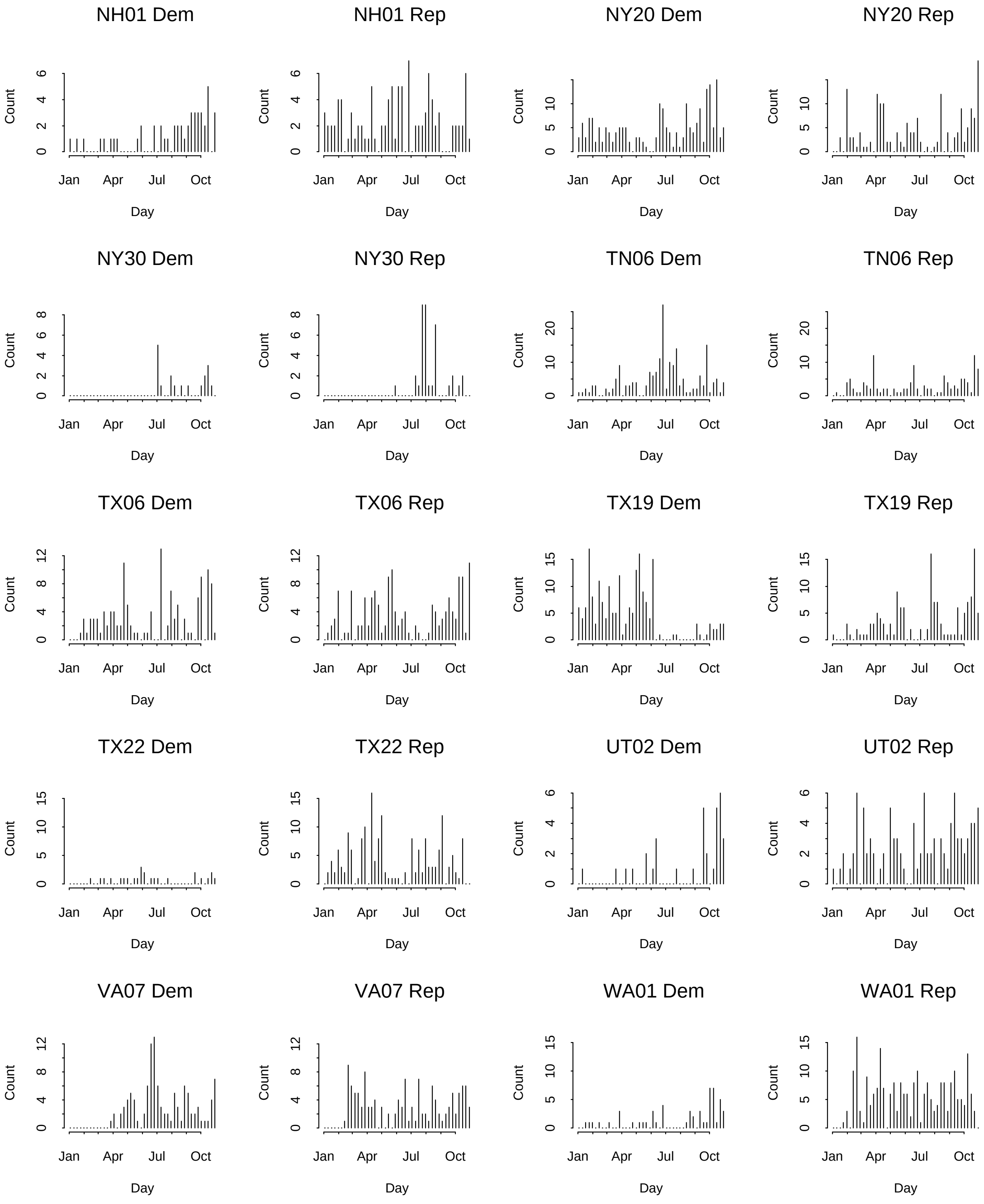

Figure 2: Weekly Counts of Clusters of Contributions to House Candidates: Individuals

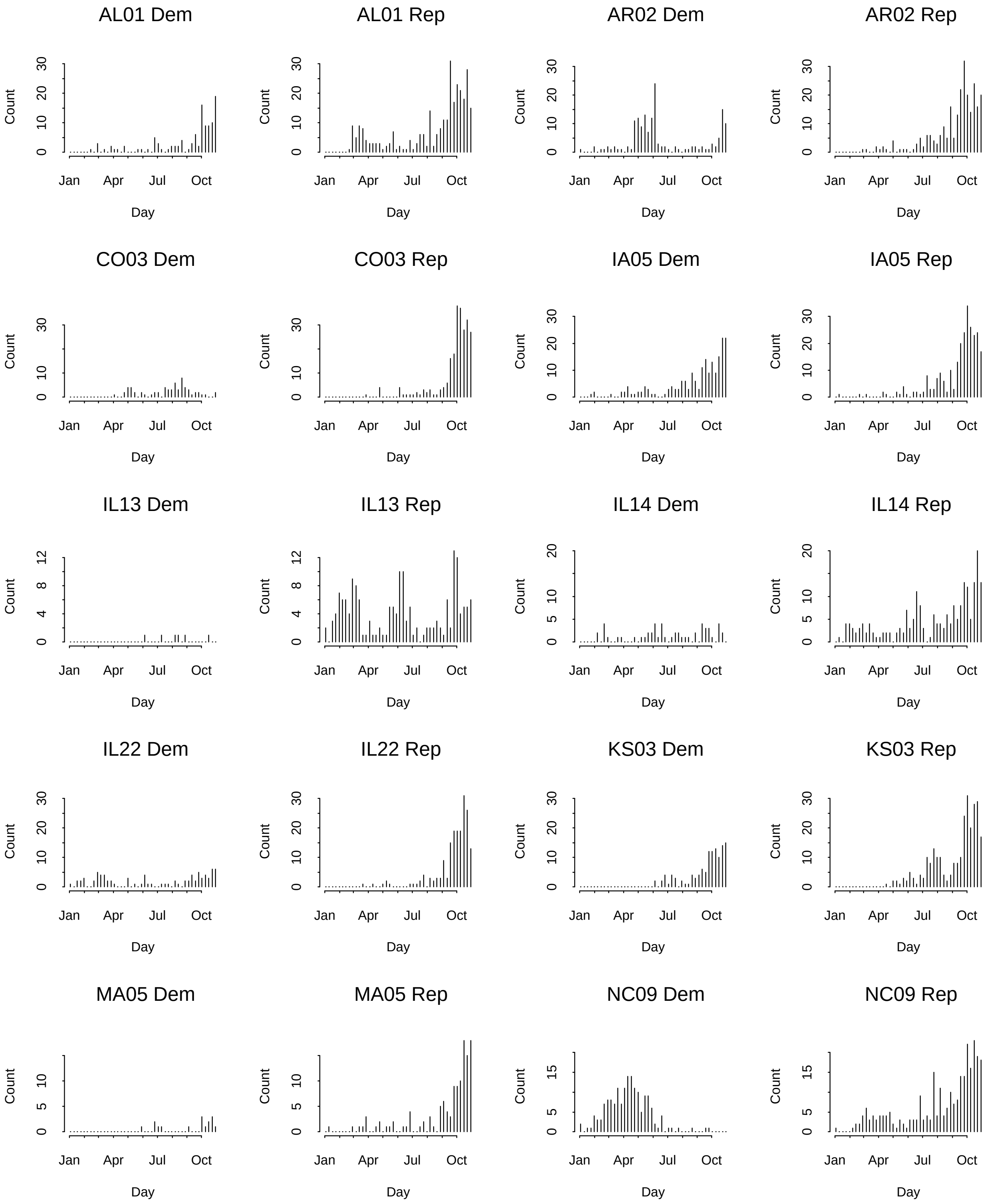

Figure 3: Weekly Counts of Clusters of Contributions to House Candidates: PACs

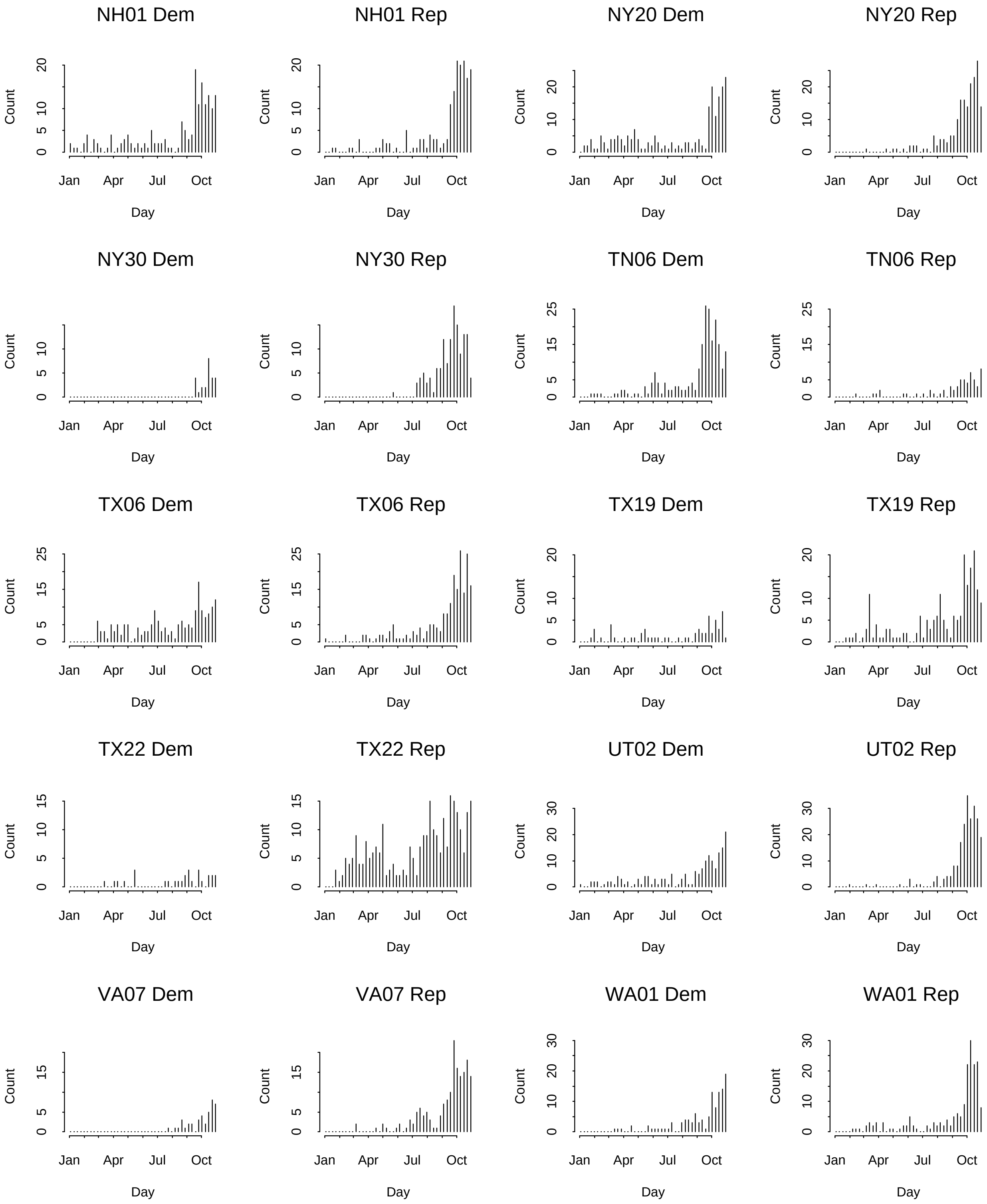

Figure 4: Weekly Counts of Clusters of Contributions to House Candidates: PACs