Supplementary Appendix A for "Lawyers as Lobbyists": Dynamic Panel Estimation of Bank Holding Company Legal Spending

Brian Libgober (Northwestern)
Daniel Carpenter (Harvard)

October 25, 2023

In this document we describe a dynamic panel procedure for estimating the effect of different variables upon bank holding company (BHC) legal spending. The document describes the data and the estimation procedures used to generate Table 5 in the paper. The basic idea is as follows. It is difficult to observe direct expenditure on legal services (including legal services that ought plausibly to be classified as lobbying) because legal service payments are often bundled. The same firm that is hired to assist with mergers and acquisitions may also be hired to fend off costly civil litigation, and the same firm may also assist with regulatory advocacy. When companies of any sort hire external agents the optimal contract is usually composed of two components - a fixed salary and a contingent payment that rises or falls in accordance with a best-response function to observables. In an important paper, [Levin 2003] shows the optimality of this arrangement as a stationary strategy for a wide range of dynamic "relational" contracts, where the constraints are imposed primarily (in some cases entirely) by the shadow of the future. In the case where companies hire lawyers externally they often engage in a dynamic contract that sets a level of pay (which may include a *retainer fee* for purchasing an option on the law firm's availability in the future) plus a contingency fee that reflects agent performance.

The idea underlying our estimates in Table 5 is that while the analyst observes only total legal spending for a BHC in a given year, she can use a large dataset with granular fixed effects for BHC and time, combined with many time-varying correlates, to estimate the marginal effect of regulatory advocacy activity in legal spending net of a wide range of factors, observed and unobserved, that would also shape legal spending. This strategy depends heavily upon the vector of controls, but the advantage of examining bank holding companies is that they are some of the most heavily regulated and documented financial entities in the world, with stiff quarterly reporting requirements for a range of assets, liabilities, income and operations. [1]

We note that our data include only *external* BHC legal expenditures and do not include yearly in-house expenditures (that is, what the BHC's own legal counsel spends on counsel that it employs directly). This would be a problem if we were trying to estimate the exact amount of legal expenditure over time, but given that we are trying to aggregate a plausible *minimum* legal expenditure attributable to regulatory advocacy it does not pose problems. Our estimates can be interpreted as the minimum legal expenditure associated with regulatory advocacy. [2] Further support for our approach comes from the fact that it is the largest BHCs (with the largest reputed in-house legal staffs) that appear to spend the most on external legal services. While the question is clearly one for future research, it is most likely the case that internal and external legal expenditure are, across firms, complements rather than substitutes.

The dataset includes measures of BHC legal expenditure by year from 2002 to 2018, and also includes a range of other measures, including (a) bank financials including assets, liabilities and interest-bearing and non-interest-bearing expenses, (b) civil and criminal actions and judgments as measured by federal court filings (measured by the Federal Judicial Center), (c) federal enforcement actions (actions and penalties) taken across four different regulatory agencies (the Federal Reserve, the Office of the Comptroller of the Currency, the Securities and Exchange Commission and the Federal Deposit Insurance Corporation), and (d) the number of pre-NPRM meetings and NPRM comments made by the BHC on ongoing federal rulemaking at the Federal Reserve.

Due to mergers and bank failures, the resulting data composes an asymmetric panel, but one which covers 798 different bank holding companies that reported at some time from 2002 to 2018. The total legal expenditure reported by bank holding companies in this data amounts to $113.9 billion dollars. This is different from popular media reporting on banks' legal costs, which include the amount of settlements that are not supposed to be included in legal costs according to FR Y-9C instructions (Griffin and Campbell 2013).

We analyze these panel data using ANOVA and two-way fixed effect regressions with lagged dependent variables (including Arellano-Bond estimation to account for problems that arise in using a lagged endogenous variable), using forms of the estimating equations below. The reason for including a lagged dependent variable in *some*

[1] For the requirements currently listed by the Federal Reserve alone (other agencies also have reporting or filing requirements, including the Treasury Department through the Office of the Comptroller of the Currency), see `https://www.federalreserve.gov/supervisionreg/afi/bhcfilings.htm` and `https://www.federalreserve.gov/apps/mdrm/`

[2] The existence of in-house legal expenditure would be a problem if the external law firm were being used as a shell for parking assets not associated with legal expenditure and later to be returned to management or shareholders, but there are laws against such asset parking as well as laws against mis-reporting of legal expenses on Federal Reserve forms, and we presume that BHCs and law firms observe these rules.

sort of specification is that the process of budgeting in many organizations may involve serial correlation of amounts and strategies [Padgett 1980].

Yet models with lagged dependent variables also have their problems. The problem with lagged dependent variables is that in expectation, a stationary series of values is correlated with the error term. The approach here is to follow the econometric literature and estimate dynamic panel models that properly specify autoregressive variables while also reducing or eliminating bias from the inclusion of a lagged dependent variable in the equation. We follow the classic approach of [Arellano and Bond 1991] as implemented by [Roodman 2009].

1 Basic Structure of the BHC Legal Spending Data and Dynamic Panel Model

To keep notation clear, we describe legal spending for company i in year t as L_{it} and express lags directly, that is the lag of L_{it} is $L_{i,t-1}$. Our panel presumes a set $\mathcal{C}$ of bank holding companies (BHCs) of size N^c indexed by i, each of which is observed for a non-zero number of years. Letting $T = \sup_{i \in \mathcal{C}} t$, then define

$$\tau_i = \sum_{t=1}^{t=T} \mathbb{1}[L_{it} \neq \emptyset]$$

The total sample size is then given by $N = \sum_{i=0}^{N^c} \tau_i$. As in general models of this sort, it is presumed that $T < N^c$, so that the panel is "wider than deep" or "small T, large N^c" ([Roodman 2009]: 128). With 798 different bank holding companies in our data and yearly data running from 2002 to 2018, these conditions are well satisfied. In particular, the data we employ exhibit the following properties.

- N^c is large, such that large sample properties apply cross-sectionally
- T is sufficiently small that large sample properties do not apply temporally
- T is sufficiently large that two- or three-step differencing does not sacrifice significant sample size
- $N^c > T$

Each of these properties, as well as their intersection, is consistent with the application of the models presented in [Arellano and Bond 1991] and [Roodman 2009], and the advisory list of Roodman (2009: 128).

Noting that $L_{i,t-1}$ has observations running from $t = 1$ to a maximum of $T - 1$, and defining the expectations operator by $\mathcal{E}$, we begin with the following fundamental estimating equation

$$L_{it} = \alpha + \delta L_{i,t-1} + \beta^{'}\mathbf{X}_{it} + \gamma^{'}R_{it} + m_t + u_{it} \quad (1)$$

where

$$u_{it} = c_i + \epsilon_{it} \quad (2)$$

and where in all forms of estimation, it is presumed that

$$\mathcal{E}(c_i) = \mathcal{E}(\epsilon_{it}) = \mathcal{E}(c_i\epsilon_{it}) = 0 \quad (3)$$

and where L measures bank legal expenses, $\mathbf{X}$ is a set of control variables varying over bank and year, R measures regulatory advocacy factors (observed meetings and rulemaking comments), β and γ are column vectors of coefficients, c_i specifies a set of bank-holding-company-specific fixed effects and m_t specifies a set of year-specific fixed effects, and ϵ represents unobservable error. The vector $\mathbf{X}$ and the variable R can also include leads and lags of relevant variables, which we do not state initially in the equation for reasons of simplicity and space, but which we report in relevant tables.

Transforming equation (1) by differencing the left-hand side we achieve

$$\Delta L_{it} = \rho\Delta L_{i,t-1} + \Delta\mathbf{X}_{it}\beta + \Delta R_{it}\gamma + \Delta\epsilon_{it} \quad (4)$$

2 Differencing and GMM estimation

The problem with both equation (1) and equation (4) is that the error terms may be correlated with the lagged dependent variables, even under differencing, as the lagged dependent variable contains $L_{i,t-1}$ and the last term in equation (4) contains $\epsilon_{i,t-1}$. There are two strategies by which $L_{i,t-1}$ can be instrumented. Before turning to instrumentation strategies, we review the basics of the GMM model (following [Roodman 2009]).

2.1 GMM Estimation

Let $\mathbf{x} = (\mathbf{x}_1, \cdots, \mathbf{x}_k)^{'}$ be a column vector of k regressors, and $\mathbf{z} = (\mathbf{z}_1, \cdots, \mathbf{z}_j)^{'}$ represents a column vector of j instruments. We allow $\mathbf{x}$ and $\mathbf{z}$ to share elements in common, and presume that $j \geq k$. Including regulatory variables R in $\mathbf{x}$ and $\mathbf{X}$ so as to observe $\mathbf{x}^R$ and $\mathbf{X}^R$, we then use $\mathbf{X}^R$, $\mathbf{L}$ and $\mathbf{Z}$ to represent matrices of N observations for $\mathbf{x}$, L and $\mathbf{z}$, and define $\mathbf{E} = \mathbf{L} - \mathbf{X}^R\hat{\beta}^R$, with the coefficient vector now subsuming γ. Given coefficient estimates of β^R, we observe residuals $\hat{\mathbf{E}} = (\hat{e}_1, \cdots, \hat{e}_N)^{'} = \mathbf{L} - \mathbf{X}^R\beta^R$. The error covariance matrix is $\mathcal{E}(\mathbf{E}\mathbf{E}^{'}) = \mathbf{\Omega}$. Letting $\mathbf{H}$ represent a candidate estimator for $\mathbf{\Omega}$, a two step estimator for β^R is

$$\hat{\beta}_1^R = (\mathbf{X^R}^{'}\mathbf{Z}(\mathbf{Z}^{'}\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}^{'}\mathbf{X^R})^{-1}\mathbf{X^R}^{'}\mathbf{Z}(\mathbf{Z}^{'}\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}^{'}\mathbf{L} \quad (5)$$

$$\hat{\beta}_2^R = (\mathbf{X^R}'\mathbf{Z}(\mathbf{Z}'\hat{\mathbf{\Omega}}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X^R})^{-1}\mathbf{X^R}'\mathbf{Z}(\mathbf{Z}'\hat{\mathbf{\Omega}}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{L} \tag{6}$$

Equations (5) and (6) create a two-step Generalized Method of Moments estimator for β^R and are "customary" for instrumental variables in the sense that GMM estimation adds the relevant terms to each step's computation equation for β^R and (unlike simple least squares estimators) iterates from the first step. It is well known that properly specified GMM estimators are, under canonical assumptions, asymptotically consistent but may have small-sample bias. The question then becomes how to construct valid $\mathbf{Z}$ and $\mathbf{\Omega}$. The first problem we refer to as *instrumentation* and the second problem we refer to as *covariance matrix estimation*.

2.2 Instrumentation by further lags of legal spending

Consider the company i for which $\tau_i = T$ is maximal. The GMM estimator of [Holtz-Eakin, Newey and Rosen 1988] uses a second lag of the spending variable, sacrificing the first observation of the dataset and starting at $t = 2$, so that the instrument matrix $\mathbf{Z}$ can be stated as follows (note the ensuing asymmetry of the matrix).

$$\mathbf{Z}_i = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ L_{i1} & 0 & \cdots & 0 \\ 0 & L_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & L_{i,T-2} \end{bmatrix}$$

The key assumption in this strategy is the moment condition

$$\mathcal{E}(\mathbf{Z}'\hat{\mathbf{E}}) = 0 \implies \sum_i L_{i,t-2}\hat{\epsilon}_{it} = 0 \ \ \forall t \geq 3 \tag{7}$$

The moment condition in equation (7) is satisfied by assumption given equation (1) and equation (4), that is, that the data generating process is modeled as AR(1). It is important, nonetheless, to check the model diagnostics for AR(2) or greater dynamics, and we do so.

2.3 Instrumentation by lags of regressors predicting lagged legal spending

As an additional check upon the problem of potential correlation between the lagged dependent variable and the error terms, one can instrument for the lagged dependent variable by using lagged regressors only. Let $\mathbf{X}_{it}^R$ be the set of regressors in equation (1) with the regulatory advocacy variables included. Then the quasi-Hatanaka strategy is achieved by instrumenting for L_{t-1} using $\mathbf{X}_{i,t-1}^R$, producing again an asymmetric version for $\mathbf{Z}$.

$$\mathbf{Z}_i = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \mathbf{X}_{i1}^R & 0 & \cdots & 0 \\ 0 & \mathbf{X}_{i2}^R & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{X}_{i,T-2}^R \end{bmatrix}$$

The original strategy here was that of [Hatanaka 1974]; see [Carpenter 1996] for an implementation in political science. Note that the criticism of recent econometrics and statistics papers (Wang and Bellemare 2019) does not apply, as we are not instrumenting for the X's here in a 2SLS fashion. The aim here is instead to instrument for the lagged dependent variable.

2.4 Estimation of covariance matrix

We follow ([Roodman 2009]: 110) in beginning with an initial estimate of $\mathbf{\Omega}^*$, the covariance matrix of the transformed errors, and then re-estimate $\mathbf{\Omega}^*$ in a second-step. Roodman programs the [Windmeijer 2005] small-sample correction for the two-step standard errors. With a two-step estimation, the Windmeijer estimator can be implemented by adding the terms `robust` and `small` to the end of the `xtabond2` command.

We conduct two sets of analyses with variants of equation (1). The first entails retrieving the estimate of γ from the equation and examining the total expenditure attributable to a regulatory variable across the dataset. For this exercise, we use annual pre-NPRM meetings with the Federal Reserve, and a one-year lead and one-year lag of this variable. We also use the following extensive battery of time-varying covariates for each BHC.

2.5 Covariates

1. **Bank Holding Company Covariates**. Annual measures of total assets, loans, commitments, interest and non-interest income, commodity investments and consulting and advising expenditures.

2. **Civil Litigation Exposure**. Annual measures of the number and stage of civil cases in federal courts involving the BHC (both at the district court and appellate levels), as well as annual data on judgments rendered (judgments or settlements). BHC involvement can be as plaintiff (`_plt`) or defendant (`_def`). Our measures separate filings (`_activity`) from cases on the docket (`_docket`).

3. **Criminal Litigation Exposure**. Annual measures of the number and stage of criminal cases in federal courts involving the BHC (both at the district court and

appellate levels), as well as annual data on judgments rendered (_fine). Our measures separate filings (_activity) from cases on the docket (_docket).

4. **Federal Enforcement Exposure**. Annual measures of the number of enforcement actions (_count) and judgment/settlement amounts (_amount) for the Securities and Exchange Commission, Federal Deposit insurance Corporation, Federal Reserve Board, and Office of the Comptroller of the Currency.

5. **Mergers and Acquisition Activity**. Annual measures of mergers and acquisitions in which the BHC is successor (_succ), predecessor (_pred) or both (_internal).

6. **Regulatory Advocacy**. Annual BHC meetings with the Federal Reserve and annual comments on Dodd-Frank rules.

3 Arellano-Bond Estimation

For purposes of display here, we show the Arellano-Bond estimates with two-step GMM used for instrumenting the lagged dependent variable, with BHC legal spending in real dollars.

Note that the Arellano-Bond test for AR(2) in first differences shows no evidence of an AR(2) pattern. (It also shows weak evidence of an AR(1) pattern, but in simple lagged dependent variables models estimated in OLS a stronger though not large pattern appears, with estimates of δ ranging between 0.1 and 0.3, which also happens to be far away from a unit root that might suggest non-stationarity in the component time series.) In addition, the Hansen and Sargan tests both show little reasons for concern. Note that ([Roodman 2009]: 129) counsels users of `xtabond2` that "because of the risks, do not take comfort in a Hansen test p-value below 0.1. View higher values, such as 0.25, as potential signs of trouble." The p-values for the Hansen and Sargan test statistics are both below 0.001.

We show here the panel data results for the period covered by Dodd-Frank (after 2009), for which we observe 480 bank-holding companies. In the output that immediately follows, we do not restrict the lag structure. We then present estimates of the main variables of interest (the BHC-specific count of meetings, led, present-year and lagged) under different lag structures, including the default recommendation of [Roodman 2009].

```
xtabond2 BHClegalspend_real l.BHClegalspend_real loans_nonUSreal   foreignbalances_real totalloansleases_real
     totalassets unusedcommits_real  noninterestincome_real  nonintexpenseother_real totinterestincome
    dataprocess_real feescomms_real consultadvise_real grosscommodities_real meeting_countlag01 meeting_count
     meeting_countlead01 equity_analyst_count criminal_litigation_activity civil_plt_litigation_activity
    civil_plt_litigation_amount civil_plt_litigation_docket civil_def_litigation_activity
    civildeflitigreal_lag01 civildeflitigreal civildeflitigreal_lead01 civildef_litigdocketlead01
    civil_def_litigation_docket civildef_litigdocketlag01  civil_plt_appeals_activity
    civil_plt_appeals_docket civil_def_appeals_activity civil_def_appeals_docket legal_and_settlement_fees
    settlement_pretax fdic_ea_count fdic_ea_amount fed_ea_count fed_ea_amount fed_bhca_count fed_ofo_count
    occ_ea_count occ_ea_amount m_and_a_pred_count m_and_a_succ_count i.date if(date > 2009), gmm(l.
    BHClegalspend_real) iv(loans_nonUSreal   foreignbalances_real totalloansleases_real  totalassets
    unusedcommits_real  noninterestincome_real  nonintexpenseother_real totinterestincome dataprocess_real
    feescomms_real consultadvise_real grosscommodities_real meeting_countlag01 meeting_count
    meeting_countlead01 equity_analyst_count criminal_litigation_activity civil_plt_litigation_activity
    civil_plt_litigation_amount civil_plt_litigation_docket civil_def_litigation_activity
    civildeflitigreal_lag01 civildeflitigreal civildeflitigreal_lead01 civildef_litigdocketlead01
```

```
    civil_def_litigation_docket civildef_litigdocketlag01  civil_plt_appeals_activity
    civil_plt_appeals_docket civil_def_appeals_activity civil_def_appeals_docket legal_and_settlement_fees
    settlement_pretax fdic_ea_count fdic_ea_amount fed_ea_count fed_ea_amount fed_bhca_count fed_ofo_count
    occ_ea_count occ_ea_amount m_and_a_pred_count m_and_a_succ_count i.date) noleveleq twostep robust small
...
Dynamic panel-data estimation, two-step difference GMM
------------------------------------------------------------------------------
Group variable: rssd                            Number of obs      =      2723
Time variable : date                            Number of groups   =       480
Number of instruments = 132                     Obs per group: min =         0
F(0, 480)     =         .                                      avg =      5.67
Prob > F      =         .                                      max =         8
-----------------------------------------------------------------------------------------------
                             |               Corrected
           BHClegalspend_real | Coefficient  std. err.      t    P>|t|     [95\% conf. interval]
-----------------------------+-----------------------------------------------------------------
           BHClegalspend_real |
                          L1. |   -.060555   .1385982    -0.44   0.662    -.3328892    .2117792
                             |
              loans_nonUSreal |   .7987692   .9098533     0.88   0.380    -.9890183    2.586557
         foreignbalances_real |   .0033176   .0019082     1.74   0.083    -.0004318    .0070671
        totalloansleases_real |   -.002226   .0028262    -0.79   0.431    -.0077792    .0033272
                  totalassets |   .0001594   .0001323     1.21   0.229    -.0001005    .0004193
           unusedcommits_real |   .0060835   .0249141     0.24   0.807    -.0428707    .0550376
       noninterestincome_real |    .013888   .0034891     3.98   0.000     .0070322    .0207438
      nonintexpenseother_real |  -.0398239   .0114408    -3.48   0.001    -.0623042   -.0173436
            totinterestincome |     .00126   .0051541     0.24   0.807    -.0088674    .0113874
             dataprocess_real |  -.0737086   .0728445    -1.01   0.312    -.2168421    .0694249
               feescomms_real |   .0246693   .0254138     0.97   0.332    -.0252668    .0746054
           consultadvise_real |   .0683278   .0476886     1.43   0.153    -.0253763     .162032
        grosscommodities_real |   8.32e-06   .0018099     0.00   0.996     -.003548    .0035647
           meeting_countlag01 |   20772.18   6097.136     3.41   0.001     8791.809    32752.56
                meeting_count |   12870.22   4377.506     2.94   0.003     4268.778    21471.66
          meeting_countlead01 |   19158.41   9343.045     2.05   0.041     800.0925    37516.74
         equity_analyst_count |  -79.86659   118.2138    -0.68   0.500    -312.1471    152.4139
civil_plt_litigation_activity |  -2549.548   3803.591    -0.67   0.503    -10023.29    4924.199
  civil_plt_litigation_amount |   .3490228    .616966     0.57   0.572    -.8632651    1.561311
  civil_plt_litigation_docket |   63.85155   252.4535     0.25   0.800    -432.1989     559.902
civil_def_litigation_activity |  -1552.677   1975.322    -0.79   0.432    -5434.025    2328.671
      civildeflitigreal_lag01 |  -.5385654   .2766581    -1.95   0.052    -1.082176    .0050451
```

```
          civildeflitigreal |   .3799863    .774399     0.49   0.624    -1.141645    1.901617
   civildeflitigreal_lead01 |  -.0799328   .6068217    -0.13   0.895    -1.272288    1.112422
  civildef_litigdocketlead01 |   610.3256    474.081     1.29   0.199    -321.2049    1541.856
civil_def_litigation_docket |  -67.69155   280.5299    -0.24   0.809    -618.9099    483.5268
   civildef_litigdocketlag01 |   526.6243   240.3958     2.19   0.029     54.26617    998.9825
  civil_plt_appeals_activity |   1785.247   15362.77     0.12   0.908    -28401.35    31971.84
  civil_def_appeals_activity |    8951.96   4227.437     2.12   0.035     645.3902    17258.53
   legal_and_settlement_fees |   107.7321   508.1634     0.21   0.832    -890.7676    1106.232
          settlement_pretax |  -114.2915   33.81149    -3.38   0.001    -180.7283   -47.85464
             fdic_ea_count |  -1808.553   2082.017    -0.87   0.385    -5899.548    2282.441
            fdic_ea_amount |  -.0225392   .0095502    -2.36   0.019    -.0413046   -.0037739
              fed_ea_count |   6166.991   10454.76     0.59   0.556    -14375.76    26709.74
             fed_ea_amount |   .0008767   .0005296     1.66   0.098    -.0001639    .0019172
            fed_bhca_count |  -12853.52   12075.18    -1.06   0.288    -36580.26    10873.21
             fed_ofo_count |   13489.43   20050.55     0.67   0.501    -25908.27    52887.13
              occ_ea_count |  -6347.993   6459.617    -0.98   0.326    -19040.61    6344.628
             occ_ea_amount |   .0020835   .0009109     2.29   0.023     .0002935    .0038734
         m_and_a_pred_count |  -9768.668   11320.33    -0.86   0.389    -32012.19    12474.86
         m_and_a_succ_count |    461.089   1735.468     0.27   0.791    -2948.964    3871.142
                            |
                       date |
                      2009  |   634.5563    2012.11     0.32   0.753    -3319.075    4588.188
                      2011  |  -2325.387   3257.049    -0.71   0.476    -8725.224     4074.45
                      2012  |  -958.8151   2746.546    -0.35   0.727    -6355.555    4437.925
                      2013  |  -3156.394   3924.471    -0.80   0.422    -10867.66    4554.873
                      2014  |  -190.7982   3904.295    -0.05   0.961    -7862.421    7480.824
                      2015  |  -2216.367   5006.383    -0.44   0.658     -12053.5    7620.767
                      2016  |  -1606.932   4994.602    -0.32   0.748    -11420.92    8207.054
                      2017  |   -6576.61   6797.166    -0.97   0.334    -19932.49    6779.266
------------------------------------------------------------------------------------------------
Instruments for first differences equation
  Standard
    D.(loans_nonUSreal foreignbalances_real totalloansleases_real totalassets
    unusedcommits_real noninterestincome_real nonintexpenseother_real
    totinterestincome dataprocess_real feescomms_real consultadvise_real
    grosscommodities_real meeting_countlag01 meeting_count meeting_countlead01
    equity_analyst_count criminal_litigation_activity
    civil_plt_litigation_activity civil_plt_litigation_amount
    civil_plt_litigation_docket civil_def_litigation_activity
    civildeflitigreal_lag01 civildeflitigreal civildeflitigreal_lead01
```

```
    cividef_litigdocketlead01 civil_def_litigation_docket
    cividef_litigdocketlag01 civil_plt_appeals_activity
    civil_plt_appeals_docket civil_def_appeals_activity
    civil_def_appeals_docket legal_and_settlement_fees settlement_pretax
    fdic_ea_count fdic_ea_amount fed_ea_count fed_ea_amount fed_bhca_count
    fed_ofo_count occ_ea_count occ_ea_amount m_and_a_pred_count
    m_and_a_succ_count 2002b.date 2003.date 2004.date 2005.date 2006.date
    2007.date 2008.date 2009.date 2010.date 2011.date 2012.date 2013.date
    2014.date 2015.date 2016.date 2017.date 2018.date)
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/16).L.BHClegalspend_real
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -1.55  Pr > z =  0.121
Arellano-Bond test for AR(2) in first differences: z =   0.60  Pr > z =  0.549
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(83)   =1922.55  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(83)   = 189.46  Prob > chi2 =  0.000
  (Robust, but weakened by many instruments.)
```

3.1 Robustness check for model without lead of comments variable

In case the one-year-lead of the comments variable is forward-associated with lagged dependent variables and/or instruments, and for robustness, we run the same model without the one-year lead of the comments variable (`meeting_countlead01`).

```
xtabond2 BHClegalspend_real l.BHClegalspend_real loans_nonUSreal   foreignbalances_real totalloansleases_real
     totalassets unusedcommits_real  noninterestincome_real  nonintexpenseother_real totinterestincome
    dataprocess_real feescomms_real consultadvise_real grosscommodities_real meeting_countlag01 meeting_count
     equity_analyst_count criminal_litigation_activity civil_plt_litigation_activity
    civil_plt_litigation_amount civil_plt_litigation_docket civil_def_litigation_activity
    civildeflitigreal_lag01 civildeflitigreal civildeflitigreal_lead01 civildef_litigdocketlead01
    civil_def_litigation_docket civildef_litigdocketlag01  civil_plt_appeals_activity
    civil_plt_appeals_docket civil_def_appeals_activity civil_def_appeals_docket legal_and_settlement_fees
    settlement_pretax fdic_ea_count fdic_ea_amount fed_ea_count fed_ea_amount fed_bhca_count fed_ofo_count
    occ_ea_count occ_ea_amount m_and_a_pred_count m_and_a_succ_count i.date if(date > 2009), gmm(l.
    BHClegalspend_real) iv(loans_nonUSreal   foreignbalances_real totalloansleases_real  totalassets
    unusedcommits_real  noninterestincome_real  nonintexpenseother_real totinterestincome dataprocess_real
    feescomms_real consultadvise_real grosscommodities_real meeting_countlag01 meeting_count
    equity_analyst_count criminal_litigation_activity civil_plt_litigation_activity
    civil_plt_litigation_amount civil_plt_litigation_docket civil_def_litigation_activity
    civildeflitigreal_lag01 civildeflitigreal civildeflitigreal_lead01 civildef_litigdocketlead01
    civil_def_litigation_docket civildef_li
tigdocketlag01  civil_plt_appeals_activity civil_plt_appeals_docket civil_def_appeals_activity
    civil_def_appeals_docket legal_and_settlement_fees settlement_pretax fdic_ea_count fdic_ea_amount
    fed_ea_count fed_ea_amount fed_bhca_count fed_ofo_count occ_ea_count occ_ea_amount m_and_a_pred_count
    m_and_a_succ_count i.date) noleveleq twostep robust small
...
Dynamic panel-data estimation, two-step difference GMM
------------------------------------------------------------------------------
Group variable: rssd                            Number of obs      =      2723
Time variable : date                            Number of groups   =       480
Number of instruments = 131                     Obs per group: min =         0
F(0, 480)     =         .                                      avg =      5.67
Prob > F      =         .                                      max =         8
-----------------------------------------------------------------------------------------------
                             |               Corrected
          BHClegalspend_real | Coefficient  std. err.      t    P>|t|     [95\% conf. interval]
-----------------------------+-----------------------------------------------------------------
          BHClegalspend_real |
```

L1.	-.0068728	.1287332	-0.05	0.957	-.2598231	.2460775
loans_nonUSreal	.8293881	1.07384	0.77	0.440	-1.280621	2.939397
foreignbalances_real	.0034521	.0020591	1.68	0.094	-.0005938	.0074981
totalloansleases_real	-.0021935	.00239	-0.92	0.359	-.0068896	.0025025
totalassets	.000114	.0001529	0.75	0.456	-.0001864	.0004144
unusedcommits_real	.0233116	.0306336	0.76	0.447	-.0368809	.0835041
noninterestincome_real	.0133563	.0040443	3.30	0.001	.0054095	.021303
nonintexpenseother_real	-.0369733	.0117957	-3.13	0.002	-.0601508	-.0137958
totinterestincome	.0047708	.0057345	0.83	0.406	-.0064971	.0160387
dataprocess_real	-.0584783	.0693182	-0.84	0.399	-.1946829	.0777262
feescomms_real	.0238176	.0258957	0.92	0.358	-.0270655	.0747006
consultadvise_real	.070264	.0526555	1.33	0.183	-.0331998	.1737278
grosscommodities_real	.0014251	.0018363	0.78	0.438	-.002183	.0050333
meeting_countlag01	18812.05	6069.216	3.10	0.002	6886.536	30737.57
meeting_count	13932.13	5043.823	2.76	0.006	4021.428	23842.83
equity_analyst_count	-11.55705	105.6097	-0.11	0.913	-219.0715	195.9574
civil_plt_litigation_activity	-2015.119	2910.404	-0.69	0.489	-7733.826	3703.587
civil_plt_litigation_amount	.1820143	.7899425	0.23	0.818	-1.370158	1.734187
civil_plt_litigation_docket	76.82058	239.1502	0.32	0.748	-393.09	546.7312
civil_def_litigation_activity	-1177.289	2097.898	-0.56	0.575	-5299.488	2944.91
civildeflitigreal_lag01	-.3380116	.3228694	-1.05	0.296	-.9724236	.2964005
civildeflitigreal	.2677789	.7416788	0.36	0.718	-1.189559	1.725117
civildeflitigreal_lead01	-.5098904	.7465646	-0.68	0.495	-1.976829	.9570482
civildef_litigdocketlead01	846.2489	552.3394	1.53	0.126	-239.053	1931.551
civil_def_litigation_docket	-300.5014	301.8723	-1.00	0.320	-893.6559	292.6532
civildef_litigdocketlag01	465.7877	227.8467	2.04	0.041	18.08741	913.4879
civil_plt_appeals_activity	1570.189	13723.66	0.11	0.909	-25395.68	28536.06
civil_def_appeals_activity	8149.191	4476.703	1.82	0.069	-647.166	16945.55
legal_and_settlement_fees	113.5017	501.4143	0.23	0.821	-871.7365	1098.74
settlement_pretax	-110.8108	36.91833	-3.00	0.003	-183.3523	-38.26933
fdic_ea_count	-2103.043	1801.343	-1.17	0.244	-5642.535	1436.448
fdic_ea_amount	-.0190653	.010386	-1.84	0.067	-.0394729	.0013422
fed_ea_count	6011.385	10569.16	0.57	0.570	-14756.15	26778.92
fed_ea_amount	.0008579	.0004712	1.82	0.069	-.0000679	.0017837
fed_bhca_count	-7676.805	9945.266	-0.77	0.441	-27218.44	11864.83
fed_ofo_count	14365.7	19676.65	0.73	0.466	-24297.3	53028.71
occ_ea_count	-4793.304	6860.61	-0.70	0.485	-18273.84	8687.235
occ_ea_amount	.0019146	.0007252	2.64	0.009	.0004897	.0033395
m_and_a_pred_count	-7395.489	11086.52	-0.67	0.505	-29179.61	14388.63

```
          m_and_a_succ_count |  -286.5335   1522.744    -0.19   0.851    -3278.601    2705.534
                             |
                        date |
                       2009  |  -1280.043    2437.19    -0.53   0.600    -6068.923    3508.836
                       2011  |  -918.6591   3317.112    -0.28   0.782    -7436.514    5599.196
                       2012  |  -2432.536    3130.32    -0.78   0.437    -8583.359    3718.286
                       2013  |  -4534.338   5767.886    -0.79   0.432    -15867.77    6799.089
                       2014  |  -3129.933   4046.001    -0.77   0.440       -11080    4820.129
                       2015  |  -4060.019   4896.698    -0.83   0.407    -13681.63    5561.593
                       2016  |  -5172.481   6621.846    -0.78   0.435    -18183.87    7838.906
                       2017  |  -8098.574   7204.415    -1.12   0.262    -22254.66    6057.514
------------------------------------------------------------------------------------------------
Instruments for first differences equation
  Standard
    D.(loans_nonUSreal foreignbalances_real totalloansleases_real totalassets
    unusedcommits_real noninterestincome_real nonintexpenseother_real
    totinterestincome dataprocess_real feescomms_real consultadvise_real
    grosscommodities_real meeting_countlag01 meeting_count
    equity_analyst_count criminal_litigation_activity
    civil_plt_litigation_activity civil_plt_litigation_amount
    civil_plt_litigation_docket civil_def_litigation_activity
    civildeflitigreal_lag01 civildeflitigreal civildeflitigreal_lead01
    civildef_litigdocketlead01 civil_def_litigation_docket
    civildef_litigdocketlag01 civil_plt_appeals_activity
    civil_plt_appeals_docket civil_def_appeals_activity
    civil_def_appeals_docket legal_and_settlement_fees settlement_pretax
    fdic_ea_count fdic_ea_amount fed_ea_count fed_ea_amount fed_bhca_count
    fed_ofo_count occ_ea_count occ_ea_amount m_and_a_pred_count
    m_and_a_succ_count 2002b.date 2003.date 2004.date 2005.date 2006.date
    2007.date 2008.date 2009.date 2010.date 2011.date 2012.date 2013.date
    2014.date 2015.date 2016.date 2017.date 2018.date)
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/16).L.BHClegalspend_real
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -1.62  Pr > z =  0.105
Arellano-Bond test for AR(2) in first differences: z =   0.72  Pr > z =  0.471
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(83)   =1984.86  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(83)   = 173.37  Prob > chi2 =  0.000
```

(Robust, but weakened by many instruments.)

3.2 Robustness check for different lag structures for Arellano-bond instrumentation

Following ([Roodman 2009]: 128-129), we report robustness checks for different lags. Note that following Roodman's counsel to "put every regressor into the instrument matrix, **Z**, in some form," we always include the full list of predictors in the instrument matrix. This places a lower bound upon the number of instruments, and constraint of instruments is then managed through lags using the STATA subfunction `laglimits`. We focus on the estimates for the meeting count variables here, as these are central to our estimation in the fully-specified models.

3.2.1 For sample covering Dodd-Frank only

Meetings coefficient estimates with lags two and longer of instruments for Arellano-Bond estimation (`laglength( 2 .)`); this is "the standard treatment for endogenous variables" ([Roodman 2009]: 124).

BHClegalspend_real	Coefficient	Corrected std. err.	t	P>|t|	[95\% conf.	interval]
meeting_countlag01	20739.31	6267.523	3.31	0.001	8424.14	33054.48
meeting_count	11584.05	4771.629	2.43	0.016	2208.189	20959.92
meeting_countlead01	17135.05	9687.64	1.77	0.078	-1900.374	36170.47

Meetings coefficient estimates with one lag of instruments for Arellano-Bond estimation (`laglength( 2 3)`)

BHClegalspend_real	Coefficient	Corrected std. err.	t	P>|t|	[95\% conf.	interval]
meeting_countlag01	15240.75	9021.958	1.69	0.092	-2486.665	32968.16
meeting_count	13783.96	7374.218	1.87	0.062	-705.7724	28273.7
meeting_countlead01	12859.82	10552.48	1.22	0.224	-7874.93	33594.58

Meetings coefficient estimates with two lags of instruments for Arellano-Bond estimation (`laglength( 2 4)`)

BHClegalspend_real	Coefficient	Corrected std. err.	t	P>|t|	[95\% conf. interval]	
meeting_countlag01	18863.85	5883.883	3.21	0.001	7302.498	30425.2
meeting_count	15568.07	4991.878	3.12	0.002	5759.432	25376.7
meeting_countlead01	16548.92	9940.353	1.66	0.097	-2983.064	36080.9

Meetings coefficient estimates with seven lags of instruments for Arellano-Bond estimation (`laglength( 2 9)`)

BHClegalspend_real	Coefficient	Corrected std. err.	t	P>|t|	[95\% conf. interval]	
meeting_countlag01	20538.61	5944.78	3.45	0.001	8857.606	32219.62
meeting_count	12523.37	4688.382	2.67	0.008	3311.082	21735.66
meeting_countlead01	16814.75	9517.893	1.77	0.078	-1887.136	35516.63

3.2.2 For sample covering from 2002 through Dodd-Frank

We now show the extended panel results. Estimates are smaller and not always statistically significant for the panel including observations before Dodd-Frank. However, the lagged and lead variables show consistency in parameter estimates, with statistically significant estimates for the standard specification (Roodman 2009).

Meetings coefficient estimates with lags two and longer of instruments for Arellano-Bond estimation (`laglength( 2 .)`); this is "the standard treatment for endogenous variables" ([Roodman 2009]: 124).

BHClegalspend_real	Coefficient	Corrected std. err.	t	P>|t|	[95\% conf.	interval]
meeting_countlag01	13526.44	6676.63	2.03	0.043	416.1808	26636.69
meeting_count	111.1697	5505.01	0.02	0.984	-10698.49	10920.83
meeting_countlead01	15702.27	8951.433	1.75	0.080	-1874.793	33279.34

Meetings coefficient estimates with one lag of instruments for Arellano-Bond estimation (`laglength( 2 3)`)

BHClegalspend_real	Coefficient	Corrected std. err.	t	P>|t|	[95\% conf.	interval]
meeting_countlag01	15563.89	7831.316	1.99	0.047	186.2877	30941.48
meeting_count	5424.361	9670.98	0.56	0.575	-13565.61	24414.33
meeting_countlead01	16246.35	10288.99	1.58	0.115	-3957.137	36449.84

Meetings coefficient estimates with two lags of instruments for Arellano-Bond estimation (`laglength( 2 4)`)

BHClegalspend_real	Coefficient	Corrected std. err.	t	P>|t|	[95\% conf. interval]	
meeting_countlag01	16526.09	8925.351	1.85	0.065	-999.765	34051.94
meeting_count	3365.494	9690.204	0.35	0.728	-15662.22	22393.21
meeting_countlead01	16468.35	10341.14	1.59	0.112	-3837.55	36774.25

Meetings coefficient estimates with seven lags of instruments for Arellano-Bond estimation (`laglength( 2 9)`)

BHClegalspend_real	Coefficient	Corrected std. err.	t	P>|t|	[95\% conf. interval]	
meeting_countlag01	13564.1	6878.4	1.97	0.049	57.65047	27070.55
meeting_count	6.408389	5534.121	0.00	0.999	-10860.41	10873.23
meeting_countlead01	15059.21	8916.802	1.69	0.092	-2449.853	32568.27

4 ANOVA for panel model

For purposes of computing estimates of variance explained by our meetings variables, we conduct ANOVA with the variables used in the Arellano-Bond models, assigning a "continuous marker (`c.*`) to each variable that takes on non-integer values. We perform two analyses here, one with only the present value and lag of the meetings variable, the other including the lead of the meetings variable, which often produces substantively and statistically significant estimates.

4.1 ANOVA for panel variables with present-value and one-year- lag of meetings.

```
anova BHClegalspend_real c.loans_nonUSreal  c.foreignbalances_real c.totalloansleases_real  c.totalassets c.
    unusedcommits_real  c.noninterestincome_real  c.nonintexpenseother_real c.totinterestincome c.
    dataprocess_real c.feescomms_real c.consultadvise_real c.grosscommodities_real c.equity_analyst_count c.
    criminal_litigation_activity c.civil_plt_litigation_activity c.civil_plt_litigation_amount c.
    civil_plt_litigation_docket c.civil_def_litigation_activity c.civildeflitigreal_lag01 c.civildeflitigreal
     c.civildeflitigreal_lead01 c.civildef_litigdocketlead01 c.civil_def_litigation_docket c.
    civildef_litigdocketlag01  c.civil_plt_appeals_activity c.civil_plt_appeals_docket c.
    civil_def_appeals_activity c.civil_def_appeals_docket c.legal_and_settlement_fees c.settlement_pretax c.
    fdic_ea_count c.fdic_ea_amount fed_ea_count c.fed_ea_amount fed_bhca_count fed_ofo_count occ_ea_count c.
    occ_ea_amount m_and_a_pred_count m_and_a_succ_count meeting_countlag 01 meeting_count i.rssd i.date

                         Number of obs =      5,988    R-squared     =  0.9518
                         Root MSE      =    38557.6    Adj R-squared =  0.9436

                  Source | Partial SS         df         MS        F    Prob>F
            ------------+----------------------------------------------------
                   Model |  1.502e+14        868   1.730e+11    116.36  0.0000
                         |
             loans_non~l |  9.391e+08          1   9.391e+08      0.63  0.4268
             foreignba~l |  3.029e+11          1   3.029e+11    203.77  0.0000
             totalloan~l |  7.544e+09          1   7.544e+09      5.07  0.0243
             totalassets |          0          0
             unusedcom~l |  8.402e+11          1   8.402e+11    565.14  0.0000
             nonintere~l |  4.174e+10          1   4.174e+10     28.07  0.0000
             nonintexp~l |  4.673e+10          1   4.673e+10     31.43  0.0000
             totintere~e |  7.822e+11          1   7.822e+11    526.13  0.0000
```

```
dataproce~l |  1.347e+11          1   1.347e+11     90.59  0.0000
feescomms~l |  5.336e+09          1   5.336e+09      3.59  0.0582
consultad~l |  1.649e+11          1   1.649e+11    110.89  0.0000
grosscomm~l |  8.069e+10          1   8.069e+10     54.28  0.0000
equity_an~t |  2.536e+09          1   2.536e+09      1.71  0.1916
criminal_.. |   61545563          1    61545563      0.04  0.8388
civil_plt.. |  5.948e+08          1   5.948e+08      0.40  0.5271
civil_pl~nt |  4.583e+09          1   4.583e+09      3.08  0.0792
civil_plt.. |  6.793e+08          1   6.793e+08      0.46  0.4991
civil_def.. |  2.213e+11          1   2.213e+11    148.85  0.0000
civi~_lag01 |  7.172e+09          1   7.172e+09      4.82  0.0281
civildefl~l |  5.876e+11          1   5.876e+11    395.27  0.0000
civ~_lead01 |  1.693e+10          1   1.693e+10     11.39  0.0007
ci~etlead01 |  1.593e+10          1   1.593e+10     10.71  0.0011
civil_def.. |  2.075e+10          1   2.075e+10     13.96  0.0002
civ~etlag01 |  6.173e+11          1   6.173e+11    415.24  0.0000
civil_plt.. |  1.122e+09          1   1.122e+09      0.75  0.3850
civil_plt.. |          0          0
civil_def.. |  2.691e+11          1   2.691e+11    180.99  0.0000
civil_def.. |          0          0
legal_and~s |  2.022e+09          1   2.022e+09      1.36  0.2436
settlemen~x |  2.168e+09          1   2.168e+09      1.46  0.2273
fdic_ea_c~t |   57570439          1    57570439      0.04  0.8440
fdic_ea_a~t |  5.657e+09          1   5.657e+09      3.81  0.0511
fed_ea_co~t |  1.063e+12         11   9.668e+10     65.03  0.0000
fed_ea_am~t |  5.823e+10          1   5.823e+10     39.16  0.0000
fed_bhca_~t |  4.625e+10          3   1.542e+10     10.37  0.0000
fed_ofo_c~t |  1.289e+09          2   6.446e+08      0.43  0.6482
occ_ea_co~t |  1.085e+12         27   4.017e+10     27.02  0.0000
occ_ea_am~t |  4.324e+09          1   4.324e+09      2.91  0.0882
m_and_a_p~t |  4.559e+10          5   9.119e+09      6.13  0.0000
m_and_a_s~t |  5.488e+10         11   4.989e+09      3.36  0.0001
meeting~g01 |  1.630e+12         19   8.577e+10     57.69  0.0000
meeting_c~t |  2.125e+12         19   1.119e+11     75.24  0.0000
       rssd |  4.556e+12        725   6.284e+09      4.23  0.0000
       date |  3.643e+10         14   2.602e+09      1.75  0.0401
            |
   Residual |  7.610e+12      5,119   1.487e+09
------------+----------------------------------------------------
      Total |  1.578e+14      5,987   2.635e+10
```

The estimate for variance explained in Table 5 (second row of ANOVA estimates) comes from the sum of the sum of squares for the meetings variables (3.76e+12) divided by the Total sum of squares.

4.2 ANOVA for panel variables with present-value, one-year- lag and one-year lead of meetings.

```
anova BHClegalspend_real c.loans_nonUSreal  c.foreignbalances_real c.totalloansleases_real  c.totalassets c.
    unusedcommits_real  c.noninterestincome_real  c.nonintexpenseother_real c.totinterestincome c.
    dataprocess_real c.feescomms_real c.consultadvise_real c.grosscommodities_real c.equity_analyst_count c.
    criminal_litigation_activity c.civil_plt_litigation_activity c.civil_plt_litigation_amount c.
    civil_plt_litigation_docket c.civil_def_litigation_activity c.civildeflitigreal_lag01 c.civildeflitigreal
     c.civildeflitigreal_lead01 c.civildef_litigdocketlead01 c.civil_def_litigation_docket c.
    civildef_litigdocketlag01  c.civil_plt_appeals_activity c.civil_plt_appeals_docket c.
    civil_def_appeals_activity c.civil_def_appeals_docket c.legal_and_settlement_fees c.settlement_pretax c.
    fdic_ea_count c.fdic_ea_amount fed_ea_count c.fed_ea_amount fed_bhca_count fed_ofo_count occ_ea_count c.
    occ_ea_amount m_and_a_pred_count m_and_a_succ_count meeting_countlag01 meeting_count meeting_countlead01
    i.rssd i.date

                         Number of obs =      5,988    R-squared     =  0.9732
                         Root MSE      =    28814.5    Adj R-squared =  0.9685

                  Source | Partial SS         df         MS        F    Prob>F
             ------------+----------------------------------------------------
                   Model |  1.535e+14        887   1.731e+11    208.48  0.0000
                         |
             loans_non~l |  1.079e+10          1   1.079e+10     13.00  0.0003
             foreignba~l |  4.652e+10          1   4.652e+10     56.03  0.0000
             totalloan~l |  1.362e+10          1   1.362e+10     16.40  0.0001
             totalassets |          0          0
             unusedcom~l |  6.996e+11          1   6.996e+11    842.58  0.0000
             nonintere~l |  3.539e+10          1   3.539e+10     42.62  0.0000
             nonintexp~l |  3.737e+10          1   3.737e+10     45.01  0.0000
             totintere~e |  6.831e+11          1   6.831e+11    822.74  0.0000
             dataproce~l |  3.277e+10          1   3.277e+10     39.47  0.0000
             feescomms~l |  3.563e+10          1   3.563e+10     42.91  0.0000
             consultad~l |  4.061e+11          1   4.061e+11    489.08  0.0000
             grosscomm~l |  2.004e+11          1   2.004e+11    241.37  0.0000
             equity_an~t |  7.448e+09          1   7.448e+09      8.97  0.0028
             criminal_.. |   13869320          1    13869320      0.02  0.8972
             civil_plt.. |  1.630e+09          1   1.630e+09      1.96  0.1613
             civil_pl~nt |  4.289e+09          1   4.289e+09      5.17  0.0231
```

```
civil_plt.. |  6623181.1          1   6623181.1      0.01  0.9288
civil_def.. |  7.315e+08          1   7.315e+08      0.88  0.3480
civi~_lag01 |  2.110e+10          1   2.110e+10     25.42  0.0000
civildefl~l |  4.744e+10          1   4.744e+10     57.14  0.0000
civ~_lead01 |  1.775e+09          1   1.775e+09      2.14  0.1437
ci~etlead01 |  3.334e+08          1   3.334e+08      0.40  0.5263
civil_def.. |  2.388e+09          1   2.388e+09      2.88  0.0900
civ~etlag01 |  2.086e+11          1   2.086e+11    251.19  0.0000
civil_plt.. |  1.039e+09          1   1.039e+09      1.25  0.2632
civil_plt.. |          0          0
civil_def.. |  1.366e+10          1   1.366e+10     16.45  0.0001
civil_def.. |          0          0
legal_and~s |  3.029e+09          1   3.029e+09      3.65  0.0562
settlemen~x |  4.562e+10          1   4.562e+10     54.94  0.0000
fdic_ea_c~t |  2.196e+08          1   2.196e+08      0.26  0.6070
fdic_ea_a~t |  3.539e+10          1   3.539e+10     42.63  0.0000
fed_ea_co~t |  6.016e+11         11   5.469e+10     65.87  0.0000
fed_ea_am~t |  2.922e+08          1   2.922e+08      0.35  0.5531
fed_bhca_~t |  2.028e+10          3   6.759e+09      8.14  0.0000
fed_ofo_c~t |  1.108e+09          2   5.538e+08      0.67  0.5133
occ_ea_co~t |  1.067e+12         27   3.952e+10     47.60  0.0000
occ_ea_am~t |  1.834e+10          1   1.834e+10     22.09  0.0000
m_and_a_p~t |  1.779e+10          5   3.557e+09      4.28  0.0007
m_and_a_s~t |  3.956e+10         11   3.597e+09      4.33  0.0000
meeting~g01 |  2.166e+12         19   1.140e+11    137.33  0.0000
meeting_c~t |  1.767e+12         19   9.300e+10    112.01  0.0000
meeting~d01 |  3.376e+12         19   1.777e+11    214.00  0.0000
       rssd |  3.114e+12        725   4.295e+09      5.17  0.0000
       date |  2.900e+10         14   2.072e+09      2.50  0.0015
            |
   Residual |  4.234e+12      5,100   8.303e+08
------------+----------------------------------------------------
      Total |  1.578e+14      5,987   2.635e+10
```

The estimate for variance explained in Table 5 (third row of ANOVA estimates) comes from the sum of the sum of squares for the meetings variables (7.31e+12) divided by the Total sum of squares.

5 Factor Analytic Methods for Regulatory Advocacy Influence upon BHC Legal Spending

The second stratagem of analysis involves collapsing the dozens of independent variable measures we have to a handful of factors by means of principal components analysis, then examining a *regulatory advocacy factor* that combines meetings and comments and examining the total legal expenditure associated with this factor across the dataset (controlling for the other factors, bank and year fixed effects and the lagged dependent variable).

Separately, outside of the dynamic panel model, we report an analysis of variance in two forms. First, we convert the different panel variables into a set of factors, and then convert the continuous factor variables to quantiles (here deciles). We then perform ANOVA, as well as ANCOVA on the un-discretized factor variables. We also calculate the change in R-squared between models with and without the regulatory advocacy factor.

5.1 Principal Component Analysis upon Covariates

Our approach to factor analysis here is confirmatory and not exploratory, that is, we pre-specify sets of variables that are akin to one another in six different categories. Our guide is conceptual, proceeding first from a set of controls characterizing BHC operations to a set of activities that are likely to predict legal spending, especially criminal and civil litigation activity as well as federal regulatory enforcement. To repeat, these categories are:

1. **Bank Holding Company Covariates**. Annual measures of total assets, loans, commitments, interest and non-interest income, commodity investments and consulting and advising expenditures.
2. **Civil Litigation Exposure**. Annual measures of the number and stage of civil cases in federal courts involving the BHC (both at the district court and appellate levels), as well as annual data on judgments rendered (judgments or settlements). BHC involvement can be as plaintiff (`_plt`) or defendant (`_def`). Our measures separate filings (`_activity`) from cases on the docket (`_docket`).
3. **Criminal Litigation Exposure**. Annual measures of the number and stage of criminal cases in federal courts involving the BHC (both at the district court and appellate levels), as well as annual data on judgments rendered (`_fine`). Our measures separate filings (`_activity`) from cases on the docket (`_docket`).
4. **Federal Enforcement Exposure**. Annual measures of the number of enforcement actions (`_count`) and judgment/settlement amounts (`_amount`) for the Securities and Exchange Commission, Federal Deposit insurance Corporation, Federal Reserve Board, and Office of the Comptroller of the Currency.

5. **Mergers and Acquisition Activity**. Annual measures of mergers and acquisitions in which the BHC is successor (`_succ`), predecessor (`_pred`) or both (`_internal`).

6. **Regulatory Advocacy**. Annual BHC meetings with Federal Reserve and annual comments on Dodd-Frank rules.

In each case we show the Stata code for principal components analysis (`pca`) and then present the "screeplot" that displays the eigenvalues of each of the possible factors (the total number of possible factors is equivalent to the total number of covariates in each PCA exercise).

5.1.1 Bank Holding Company Covariates

```
pca loans_nonUSreal   foreignbalances_real totalloansleases_real  totalassets
     unusedcommits_real  noninterestincome_real  nonintexpenseother_real
    totinterestincome dataprocess_real feescomms_real consultadvise_real
    grosscommodities_real
screeplot, yline(1) ci(het) scheme(s2mono) graphregion(fcolor(white))
predict bhc_f1 bhc_f2, score
```

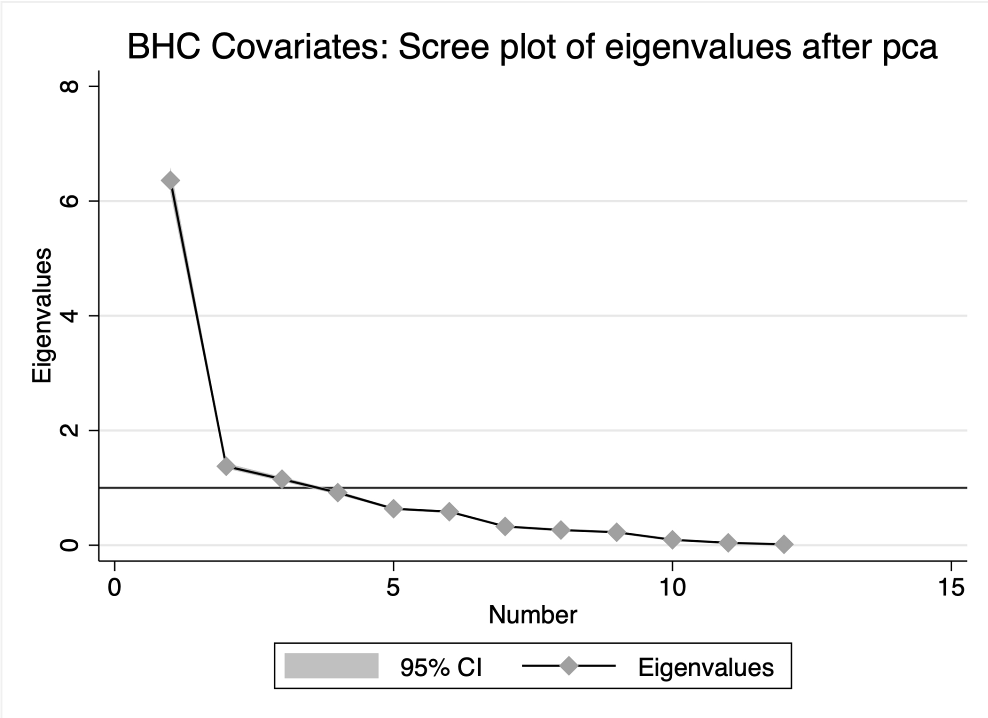

Figure 1: Eigenvalue plot for BHC covariates.

5.1.2 Civil Litigation Activity

```
pca civil_plt_litigation_activity civil_plt_litigation_amount
    civil_plt_litigation_docket civil_def_litigation_activity
    civil_def_litigation_amount civil_def_litigation_docket
    civil_plt_appeals_activity civil_plt_appeals_docket
    civil_def_appeals_activity civil_def_appeals_docket
screeplot, ci(het) yline(1) scheme(s2mono) graphregion(fcolor(white))
predict civ_f1 civ_f2, score
```

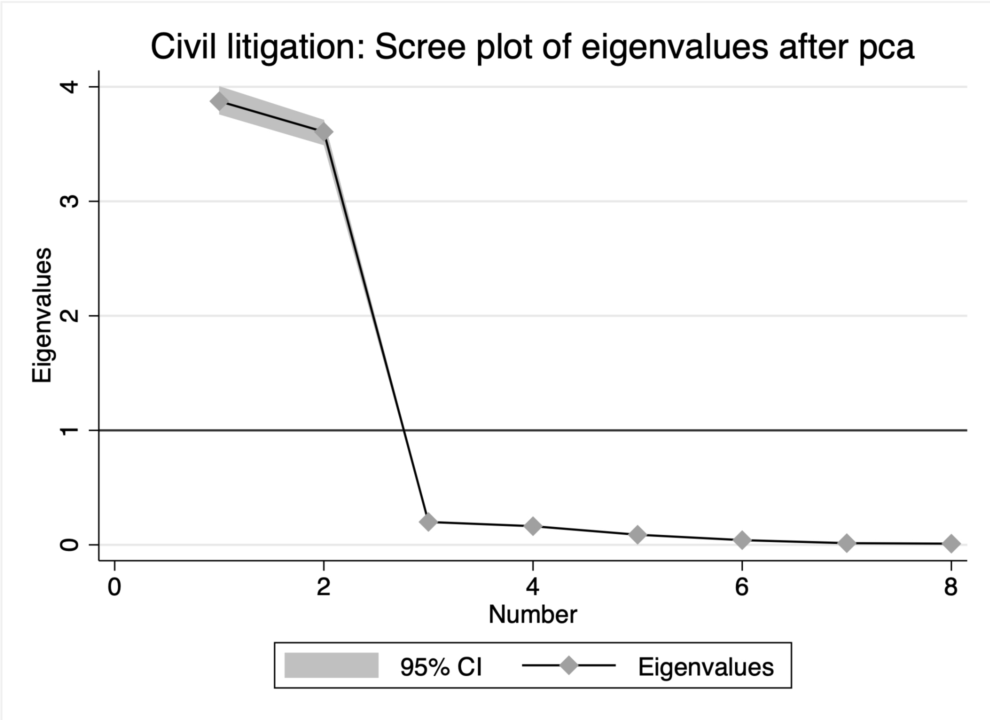

Figure 2: Eigenvalue plot for Civil Litigation covariates.

5.1.3 Criminal Litigation Exposure

```
pca criminal_litigation_activity criminal_litigation_fine
    criminal_litigation_docket criminal_appeals_activity
    criminal_appeals_docket
screeplot, yline(1) ci(het) scheme(s2mono) graphregion(fcolor(white))
predict crim_f1, score
```

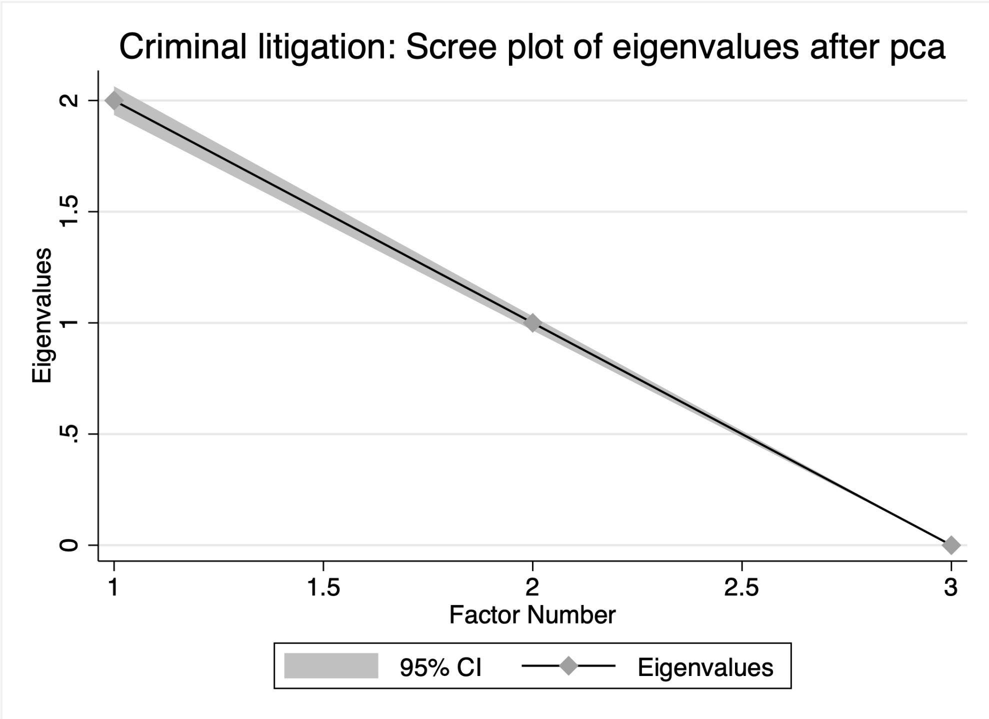

Figure 3: Eigenvalue plot for Criminal Litigation covariates.

5.1.4 Federal Regulatory Enforcement

```
pca sec_ea_count fdic_ea_count fdic_ea_amount fed_ea_count fed_ea_amount
    fed_bhca_count fed_ofo_count occ_ea_count occ_ea_amount
screeplot, yline(1) ci(het) scheme(s2mono) graphregion(fcolor(white))
predict enf_f1 enf_f2, score
```

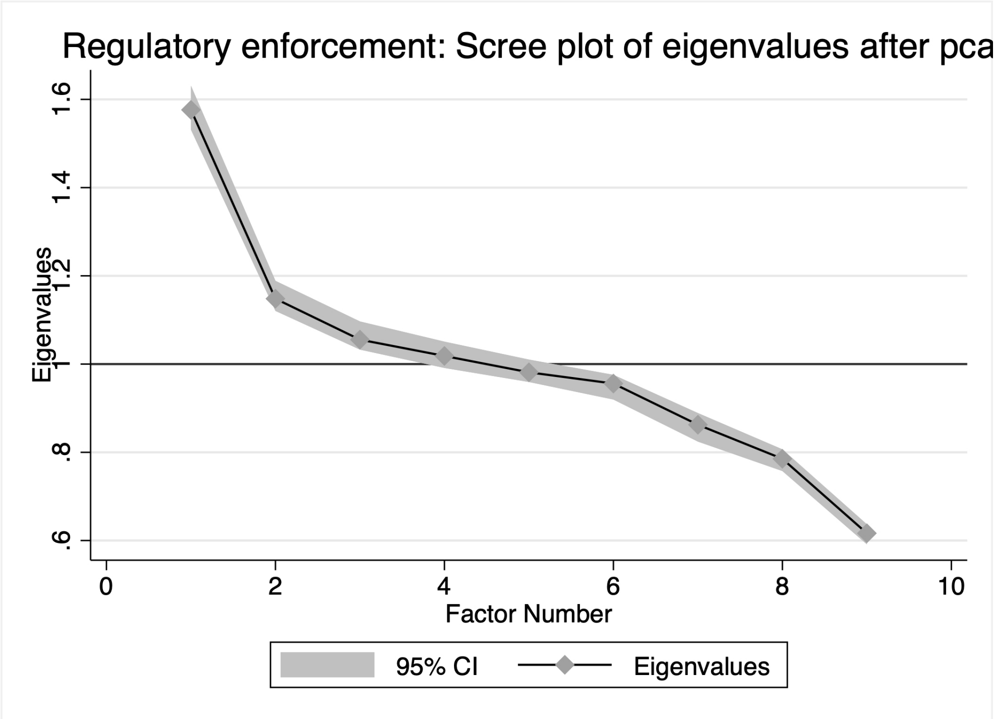

Figure 4: Eigenvalue plot for Federal Regulatory Enforcement covariates.

5.1.5 Mergers and Acquisitions Activity

```
pca m_and_a_pred_count m_and_a_succ_count m_and_a_internal_count
screeplot, yline(1) ci(het) scheme(s2mono) graphregion(fcolor(white))
predict MnA_f1, score
```

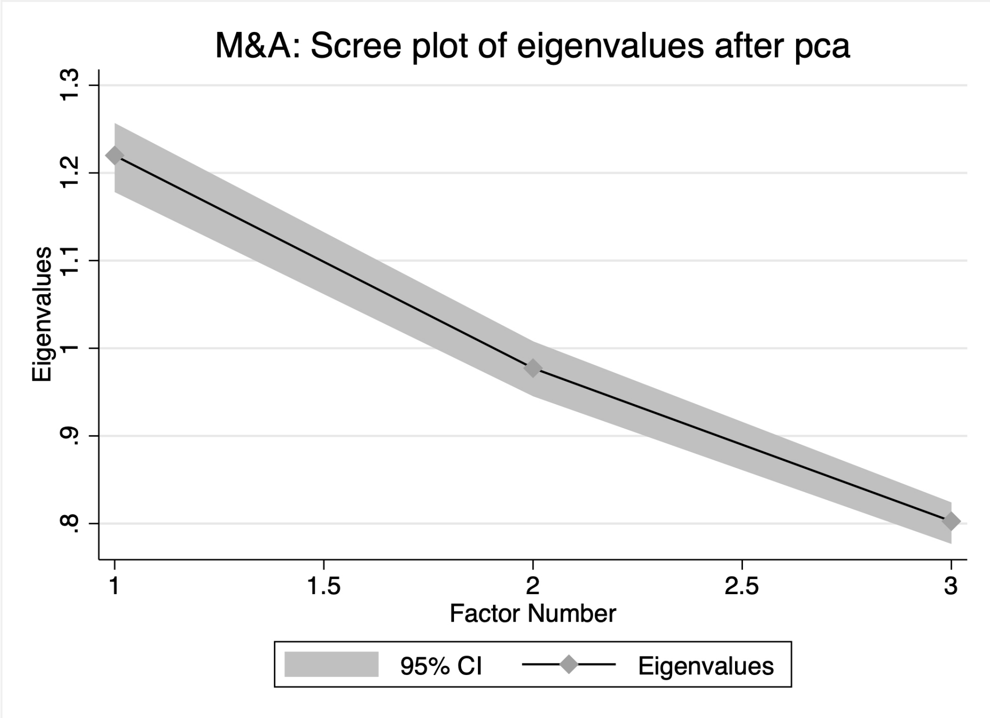

Figure 5: Eigenvalue plot for Mergers and Acquisitions Activity.

5.1.6 Regulatory Advocacy

```
pca meeting_count comment_count
screeplot, yline(1) ci(het) scheme(s2mono) graphregion(fcolor(white))
predict reg_f1, score
```

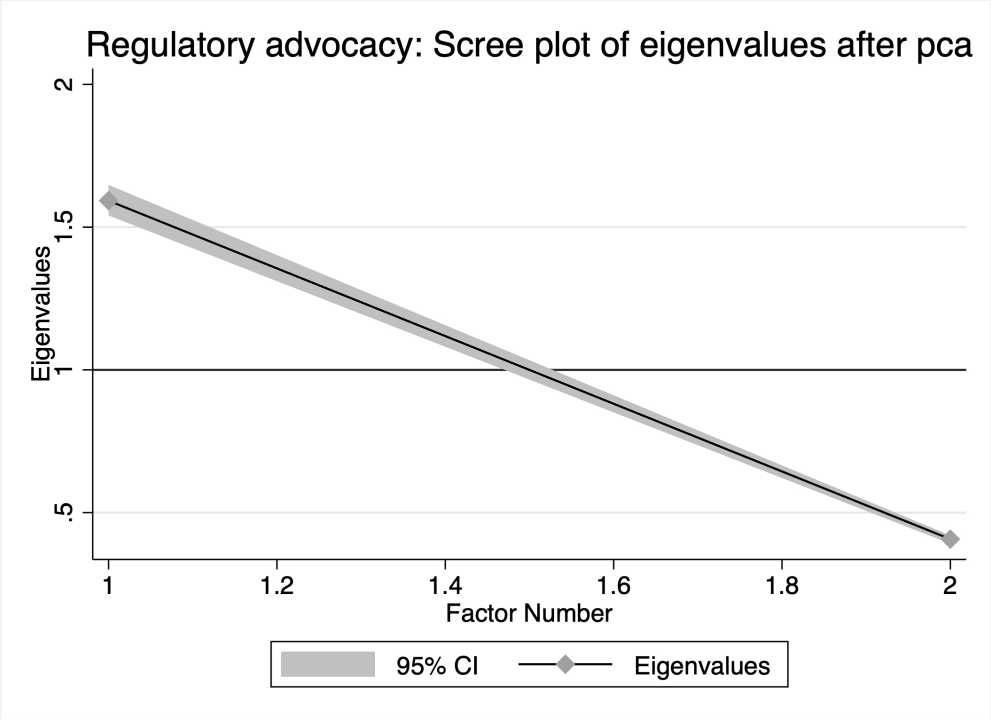

Figure 6: Eigenvalue plot for Regulatory Advocacy covariates.

5.2 ANOVA with PCA Factors

For ANOVA, we retrieve a maximum of two factors for each set of variables, neglecting a second factor for any conceptual category and associated set of variables where the associated eigenvalue is indistinguishable from unity. We then conduct a basic ANOVA with two-way fixed effects (identifying continuous variables with the `c.*` marker) producing the following results.

```
anova BHClegalspend_real c.bhc_f1 c.bhc_f2 c.civ_f1 c.civ_f2  c.enf_f1 c.enf_f2 c.MnA_f1 c.reg_f1 i.rssd i.
    date

                         Number of obs =      7,563    R-squared     =  0.7198
                         Root MSE      =    84737.1    Adj R-squared =  0.6857

                  Source | Partial SS         df         MS        F    Prob>F
              -----------+----------------------------------------------------
                   Model |  1.243e+14        821   1.515e+11     21.09  0.0000
                         |
                  bhc_f1 |  1.046e+13          1   1.046e+13   1457.25  0.0000
                  bhc_f2 |  7.510e+11          1   7.510e+11    104.59  0.0000
                  civ_f1 |  2.653e+12          1   2.653e+12    369.52  0.0000
                  civ_f2 |  2.939e+12          1   2.939e+12    409.31  0.0000
                  enf_f1 |  5.671e+11          1   5.671e+11     78.97  0.0000
                  enf_f2 |  3.915e+11          1   3.915e+11     54.52  0.0000
                  MnA_f1 |  5.361e+10          1   5.361e+10      7.47  0.0063
                  reg_f1 |  1.579e+12          1   1.579e+12    219.89  0.0000
                    rssd |  1.818e+13        797   2.281e+10      3.18  0.0000
                    date |  3.588e+11         16   2.242e+10      3.12  0.0000
                         |
                Residual |  4.840e+13      6,741   7.180e+09
              -----------+----------------------------------------------------
                   Total |  1.727e+14      7,562   2.284e+10
```

The estimate of variance explained by regulatory factors presented in Table 5 of the paper (first row of ANOVA estimates) is given by $\frac{1.579e+12}{1.727e+14}$.

It is noteworthy that both criminal litigation exposure and mergers and acquisitions activity are, in panel models or in factor analytic models, poor predictors of BHC legal spending. We have identified this as a research agenda for a different paper.

5.3 Two-Way Fixed Effects Regression with Factor Variables

We then can use the factor-analytic variables as predictors for legal spending directly in regressions. An important point to keep in mind is that the factor-analytic variables are drawn from other estimating equations and have known error. We therefore engage in randomization inference testing . We begin with the unadjusted equation 1, restated in a single line as follows.

$$L_{it} = \alpha + \delta L_{i,t-1} + \beta^{'}\mathbf{X}_{it} + \gamma^{'} R_{it} + c_i + m_t + \epsilon_{it} \tag{8}$$

We first estimate the two-way FE model with all factors.

```
xtreg BHClegalspend_real  BHClegalspendreal_lag01  bhc_f1 bhc_f2 civ_f1 civ_f2 crim_f1 enf_f1 enf_f2 MnA_f1
    reg_f1 i.date, fe i(rssd) cluster(rssd)

Fixed-effects (within) regression               Number of obs     =      6,765
Group variable: rssd                            Number of groups  =        777

R-squared:                                      Obs per group:
     Within  = 0.6468                                         min =          1
     Between = 0.7723                                         avg =        8.7
     Overall = 0.6877                                         max =         16

                                                F(24,776)         =          .
corr(u_i, Xb) = -0.7794                         Prob > F          =          .

                                            (Std. err. adjusted for 777 clusters in rssd)
-----------------------------------------------------------------------------------------
                        |               Robust
     BHClegalspend_real | Coefficient  std. err.      t    P>|t|     [95\% conf. interval]
------------------------+----------------------------------------------------------------
BHClegalspendreal_lag01 |   .3449738   .0485263     7.11   0.000     .2497154    .4402321
                 bhc_f1 |   47162.89   12312.44     3.83   0.000     22993.25    71332.53
                 bhc_f2 |  -4567.375   10888.82    -0.42   0.675    -25942.41    16807.66
                 civ_f1 |   10836.02   4777.185     2.27   0.024     1458.281    20213.76
```

```
             civ_f2 |   11603.88   4954.749     2.34   0.019     1877.581    21330.18
            crim_f1 |  -21.15025   15.02376    -1.41   0.160    -50.64227    8.341775
             enf_f1 |   11126.58   6154.641     1.81   0.071    -955.1383     23208.3
             enf_f2 |  -4476.491   1678.176    -2.67   0.008    -7770.793   -1182.189
             MnA_f1 |   1982.637   2746.954     0.72   0.471    -3409.705    7374.979
             reg_f1 |   8327.867   2511.938     3.32   0.001     3396.868    13258.87
                    |
               date |
              2004  |   2645.405   2052.687     1.29   0.198    -1384.071    6674.882
              ...
              2018  |  -23628.42   8050.733    -2.93   0.003    -39432.21   -7824.623
                    |
              _cons |   9588.532   1852.101     5.18   0.000      5952.81    13224.25
--------------------+----------------------------------------------------------------
            sigma_u |  78500.346
            sigma_e |  64103.184
                rho |  .59994069   (fraction of variance due to u_i)
-------------------------------------------------------------------------------------
```

5.4 Randomization Inference for the Regulatory Advocacy Factor

We then conduct a randomization inference exercise for the by using the `ritest` command package in Stata. The `ritest` command we use is specified for panel data in that it replicates the exact two-way fixed effects regression just estimated and uses the company indicators as strata. We present the command and basic output, followed by a kernel density plot for the recovered 9augmented) distribution.

```
ritest reg_f1 _b[reg_f1], reps(1000) strata(rssd) kdensityplot: xtreg BHClegalspend_real
    BHClegalspendreal_lag01  bhc_ f1 bhc_f2 civ_f1 civ_f2 enf_f1 reg_f1 i.date, fe i(rssd) cluster(rssd)

Resampling replications (1,000)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
...
.................................................. 1,000

      Command: xtreg BHClegalspend_real BHClegalspendreal_lag01 bhc_f1 bhc_f2 civ_f1 civ_f2 enf_f1 reg_f1 i.
          date, fe
                  i(rssd) cluster(rssd)
        _pm_1: _b[reg_f1]
  res. var(s):  reg_f1
   Resampling:  Permuting reg_f1
Clust. var(s):  __000000
     Clusters:  7563
Strata var(s):  rssd
       Strata:  798

------------------------------------------------------------------------------
T            |     T(obs)       c       n   p=c/n   SE(p) [95\% Conf. Interval]
-------------+----------------------------------------------------------------
       _pm_1 |   8818.983      44    1000  0.0440  0.0065  .0321495   .0586204
------------------------------------------------------------------------------
Note: Confidence interval is with respect to p=c/n.
Note: c = #{|T| >= |T(obs)|}
```

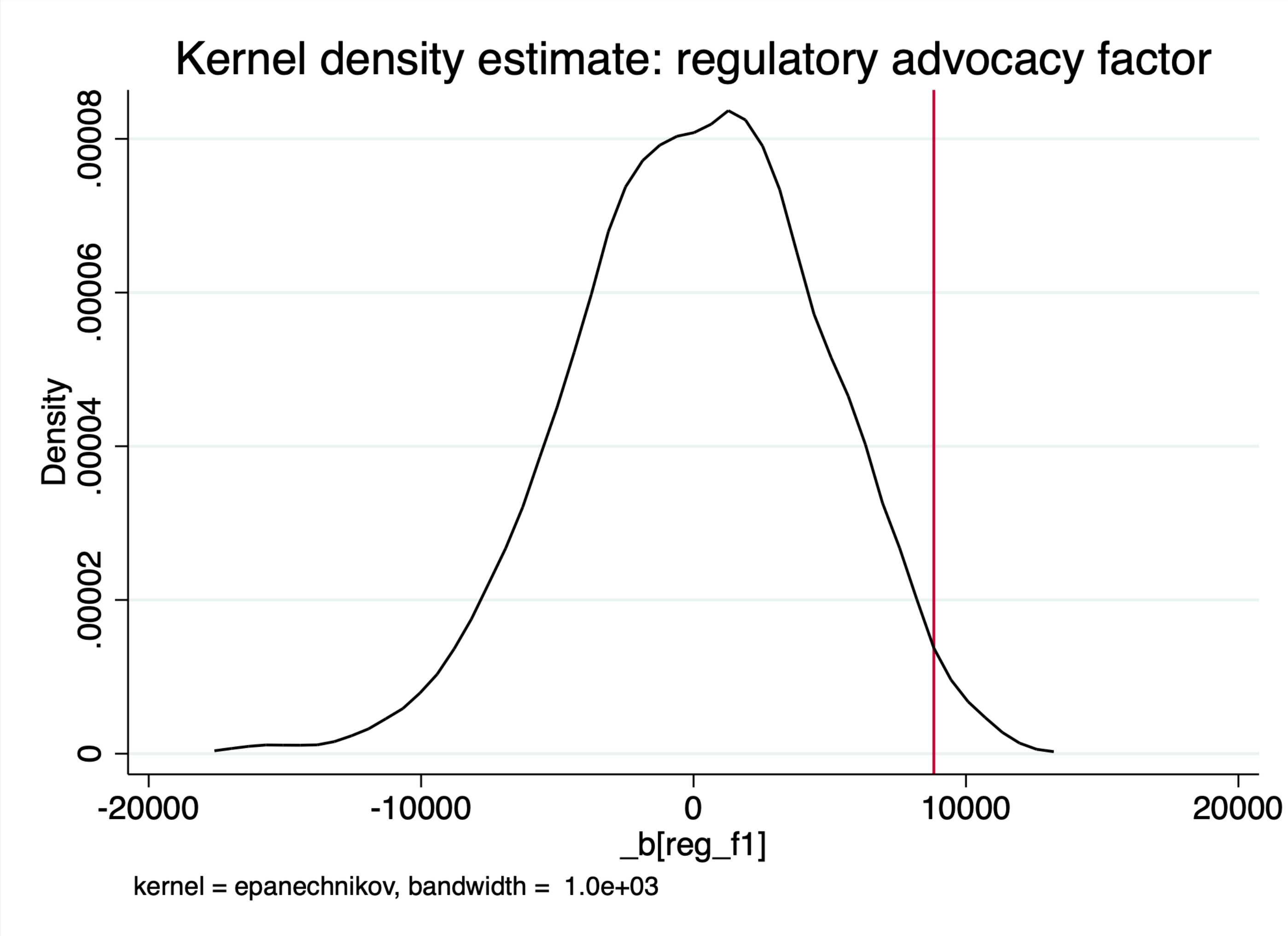

Figure 7: Kernel density plot after randomization inference (1000 resampling replications, 798 strata (BHCs).

5.5 Errors in Variables for Regulatory Advocacy

We can also address issues of errors-in-variables directly by examining the difference between a simple regression with the factor and a lagged dependent variable and the regulatory advocacy factor and an errors-in-variables regression (using `eivreg` in Stata) that augments the covariance matrix to account for the measurement error in regressors. The reliability estimate for the regulatory advocacy factor is taken from the cumulative loading of the first factor (0.796) in the `pca` output. Note that the equations are purely linear, lack fixed effects and are thus mis-specified in a strict sense. But the comparison between the unadjusted regression and the errors-in-variables regression is instructive, especially as it displays a coefficient estimate for the regulatory advocacy factor similar to the estimates retrieved for the meetings variables for the Arellano-Bond dynamic panel models.

```
reg BHClegalspend_real  BHClegalspendreal_lag01 reg_f1

      Source |       SS           df       MS      Number of obs   =     6,765
-------------+----------------------------------   F(2, 6762)      =   8812.51
       Model |  1.1426e+14         2  5.7130e+13   Prob > F        =    0.0000
    Residual |  4.3837e+13     6,762  6.4828e+09   R-squared       =    0.7227
-------------+----------------------------------   Adj R-squared   =    0.7226
       Total |  1.5810e+14     6,764  2.3373e+10   Root MSE        =     80516

------------------------------------------------------------------------------------------
     BHClegalspend_real | Coefficient  Std. err.      t    P>|t|     [95\% conf. interval]
------------------------+-----------------------------------------------------------------
BHClegalspendreal_lag01 |   .7373402    .007161   102.97   0.000     .7233024    .7513781
                 reg_f1 |   15950.68   859.1897    18.56   0.000     14266.39    17634.96
                  _cons |   3486.455   986.0444     3.54   0.000     1553.498    5419.413
------------------------------------------------------------------------------------------

. eivreg BHClegalspend_real  BHClegalspendreal_lag01  reg_f1, reliab(BHClegalspendreal_lag01 1 reg_f1 .796)

Errors-in-variables regression

                     Assumed
    Variable     reliability
----------------------------                   Number of obs     =      6,765
BHCle~_lag01          1.0000                   F(  2,      6762) =      98.48
      reg_f1          0.7960                   Prob > F          =     0.0000
```

```
          *          1.0000                   R-squared         =     0.7282
                                              Root MSE          =    79720.7

------------------------------------------------------------------------------------------
     BHClegalspend_real | Coefficient  Std. err.      t    P>|t|     [95\% conf. interval]
------------------------+-----------------------------------------------------------------
BHClegalspendreal_lag01 |    .710863   .0936423     7.59   0.000     .5272946    .8944314
                 reg_f1 |    22100.2    9877.07     2.24   0.025     2738.035    41462.37
                  _cons |   3845.217    1210.99     3.18   0.002     1471.296    6219.137
------------------------------------------------------------------------------------------
```

References

[Arellano and Bond 1991] Arellano, Manuel, and Stephen Bond. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The review of economic studies* 58, no. 2 (1991): 277-297.

[Carpenter 1996] Carpenter, Daniel P. "Adaptive signal processing, hierarchy, and budgetary control in federal regulation," *American Political Science Review* 90, no. 2 (1996): 283-302.

[Hatanaka 1974] Hatanaka, Michio. "An efficient two-step estimator for the dynamic adjustment model with autoregressive errors," *Journal of Econometrics* 2, no. 3 (1974): 199-220.

[Holtz-Eakin, Newey and Rosen 1988] Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen. "Estimating vector autoregressions with panel data," *Econometrica* (1988): 1371-1395.

[Levin 2003] Levin, Jonathan. "Relational incentive contracts," *American Economic Review* 93, no. 3 (2003): 835-857.

[Padgett 1980] Padgett, John. "Bounded Rationality in Budgetary Research." *(*American Political Science Review 74, no. 2 (1980): 354-372.

[Roodman 2009] Roodman, David. "How to do xtabond2: An introduction to difference and system GMM in Stata." *(*The stata journal 9, no. 1 (2009): 86-136.

[Windmeijer 2005] Windmeijer, Frank. "A finite sample correction for the variance of linear efficient two-step GMM estimators," *Journal of Econometrics* 126, no. 1 (2005): 25-51.