Puppetmasters or Pawns?
The Power of Congressional Staffers

PRELIMINARY WORK. DO NOT DISTRIBUTE.

Andrew Kao
Sara Ji*

September 4, 2024

Abstract

Members of Congress do not write their own bills—their staffers do. We estimate the importance of personal staffers relative to the members of Congress that they work for in the US House of Representatives. To do so, we develop an extension of the AKM mover design to team settings where an individual's output is unobserved, and estimate the importance of staffers using both frequentist and Bayesian methods. We consistently find that staffers account for 40% or more of the difference in legislative productivity between high and low performing offices. Conventional estimation methods are either infeasible, or misspecified and deliver estimates less than half as large. We also find that the production function between the team of staffers and the Representative is close to Cobb-Douglas, that staffers moderate ideological extremism, and that staffers have the greatest impact when they focus on issues that are less politically divisive.

Keywords: Congress, staffers, legislation, teams
JEL Classification: D72, D73, J44

*Kao: Harvard University. Ji: Harvard University. Email: andrewkao@g.harvard.edu.

1 Introduction

Congress produces over 10,000 bills a year, passes over 500 of these into law, and controls a 5 trillion dollar budget. Although the media typically highlights the importance of Representatives and Senators in the production and passage of bills, their staffers are the ones who draft the actual text of these bills (Fox and Hammond, 1977). How important are these staffers?

Some scholars have argued that staffers are no more than "extensions" of their member of Congress (Kingdon, 1989; DeGregorio, 1988), who are "bound" to them (Salisbury and Shepsle, 1981; Hall and Deardorff, 2006), and whose value after leaving Congress is the connection to the specific member they worked for (Blanes i Vidal et al., 2012; Bertrand et al., 2014)—a theory of staffers as "pawns." Others have argued that staffers can act autonomously (Romzek and Utter, 1997), shape policy and originate legislation (Hagedorn, 2015; Crosson et al., 2020), and influence legislative productivity (Ommundsen, 2023)—a theory of staffers as "puppetmasters." In this paper, we quantitatively estimate the importance of personal staffers relative to the members of Congress that they work for in the US House of Representatives.

Although some Congressional staffers eventually achieve a degree of fame (for instance, President Lyndon B. Johnson was a Congressional staffer from 1931 to 1937), the vast majority of staffers work away from the limelight. Their names do not appear on the bills that they write, making it difficult to quantitatively assess their importance to the legislative process. We make progress on this question by extending the canonical value added model (or mover design, Abowd et al. (1999)) to team settings where the contribution of each individual is unobserved. With this team-based mover model, we are able to estimate the extent to which legislatively productive offices are productive because of their staffers, as opposed to their Representatives. Specifically, we estimate the share of differences between above and below median output offices that is attributable to the team of staffers. We develop two novel strategies to estimate the model dynamically, allowing us to estimate the staffer share in an event study, as well as test for pre-trends and model misspecification. We also demonstrate how to estimate this team-based model using Bayesian methods, allowing us to flexibly extend the model and study a number of related questions of interest. Our Bayesian approach is suitable for the estimation of nonlinear models in teams larger than 2, a problem that the prior literature has had difficulty tackling. This estimation framework can be applied to any setting featuring team production where individual output is unobserved.

We focus our attention on personal staffers to Representatives of the House. Our core

finding is that these staffers account for 40% or more of the share of differences in legislative productivity between above and below median offices. Using a simple linear decomposition, we estimate that the staffer share is roughly 50% for the production of both bills and laws. When instrumenting for staffer movements using election turnovers and leave-out migration flows, we estimate an even larger staffer share. We then estimate dynamic effects, beginning by estimating the staffer share in an event study framework. We find no evidence of pre-trends and estimate a staffer share of 40%. An alternative exercise shows that we can dynamically predict when staffers move using nothing but their out-of-sample fixed effect; this validates our estimation procedure. We also find minimal contamination or spillovers from other staffers or the Representative. Finally, we estimate the staffer share using Bayesian methods. We once again estimate a staffer share of 40%. Including non-movers back into the design only increases the staffer share. These results collectively demonstrate the importance of staffers to the legislative process.

The Bayesian approach also allows us to estimate a variety of model extensions. We begin by testing the extent to which different jobs facilitate the production of bills. Unsurprisingly, we find that legislative-facing jobs facilitate the production of more bills. More surprisingly, we find that Chiefs of Staff are uniquely incapable of converting their inherent skill into legislation—they would contribute more to legislative output even in non-legislation-facing jobs. We then estimate the return to experience in Congress and find that experience on the Hill increases the legislative productivity of Representatives but not staffers, a finding consistent with the literature. We also find small but significant gains from diversity of all kinds: ethnic, gender, educational, and prior work experience. Next, we estimate a nested CES production function to uncover the relationship between staffers and Representatives. We find that while staffers within an office are highly substitutable for one another, Representatives and their teams of staffers have a production function that is much closer to Cobb-Douglas. This production function yields a simple test of potentially Pareto improving trades of staffers between offices. We find that there are very few such Pareto-improving trades in the data, implying that the majority of offices hire staffers in a sophisticated manner that accounts for both team and Representative skills. Finally, we examine the role that staffers play in shaping the ideological content of bills. Along the first DW-Nominate dimension (the conventional left-right ideological axis), we find that Representatives account for nearly all of the variation, with staffers playing a small role in moderating extreme offices. However, along the second DW-Nominate dimension (capturing less partisan policy issues), staffers account for a much more substantial 80% of the share in ideological differences. These results suggest ways in which staffers new to Congress can quickly make an impact: by getting into leg-

islative positions, working for Representatives that have high intrinsic productivity, and by focusing on policy issues outside of conventional partisan divides.

This paper directly contributes to three strands of literature. First, on the estimation of individual effects in team environments. To the best of our knowledge, Chan (2021), Bonhomme (2021), Bergeron et al. (2022), and Ahmadpoor and Jones (2019) are the four papers that tackle this question. None estimate a nonlinear model featuring an average team size greater than 2. We provide a framework that can efficiently estimate individual effects for arbitrarily sized teams, and utilize it in the context of Congressional House personal staffs, where teams can contain up to 18 staffers. Other studies, particularly in healthcare, have estimated individual effects in team environments without estimating the effect of an entire team: in Chen (2021), the analysis focuses only on estimating the effects of the 'main' physician on a team, while in Silver (2021), physicians work individually on cases but may be influenced by other nearby physicians via peer effects.[1] We extend on a large literature that uses the mover design in value-added models (Abowd et al., 1999),[2] to settings where production occurs in teams and individual contributions are unobserved.

Second, we contribute to the literature on the effect that Congressional staffers have on the legislative process. In economics, attention has primarily focused on the "revolving door" phenomenon and the impact of staffers-turned-lobbyists (Blanes i Vidal et al., 2012; Bertrand et al., 2014).[3] Political scientists have long debated the role that staffers play in Congress. Some have claimed that they are little more than extensions of their member of Congress,[4] while others have argued for their ability to independently influence the legislative process.[5] While this work has been largely qualitative, recent work has focused on documenting the role of staffers through controlled correlations: Leal and Hess (2004) finds that House freshmen are more likely to hire experienced staffers, while Crosson et al. (2020) finds that experienced staff can help new members of Congress and committee chairs, but not the average member of Congress. Ommundsen (2023) finds that legislative output is higher in committees with experienced staff. Although we touch on the effect of staffer experience on office productivity in Section 6.2, our primary contribution is to estimate the individual effect of staffers and decompose the differences in legisla-

[1] Also related are Freund (2024),Iranzo et al. (2008), and Weidmann et al. (2024), who directly estimate individual effects and take them into team environments.

[2] See Finkelstein et al. (2021); Abowd et al. (2008); Card et al. (2013); Bender et al. (2018); Song et al. (2019); Cantoni and Pons (2022); Chetty and Hendren (2018), among many others.

[3] Other studies on the revolving door include Shepherd and You (2020) and Lazarus et al. (2019).

[4] See, among others, Kingdon (1989); DeGregorio (1988); Salisbury and Shepsle (1981); Hall and Deardorff (2006)

[5] See, among others, Fox and Hammond (1977); Romzek and Utter (1997); Hagedorn (2015).

tive output between Congressional offices into the effect of staffers against the effect of Representatives. Closest to us in this regard is Montgomery and Nyhan (2017), who find that Congress members who exchange important staffers are more similar in legislative effectiveness and ideology, though they do not propose an identification strategy. In addition to providing a design to causally identify staffer effects, we are also able to speak to mechanisms that make certain staffers more effective than others.

Third, this study contributes to the broader question of how political appointees and bureaucrats influence the government and economy. A central tension in democracies is that while politicians are elected, appointees and bureaucrats are not, but they may nonetheless shape policy (Aberbach et al., 1981). Political appointees have been found to be influenced by partisan cycles (Spenkuch et al., Forthcoming), and efforts to reduce the partisan influence on bureaucrats have increased their productivity (Aneja and Xu, Forthcoming). We contribute to this literature by documenting the extent to which staffers may mitigate partisan divides and contribute particularly to less-partisan legislation. A rich literature has studied the structure and incentives of bureaucracies theoretically (Niskanen, 1968; Besley and Persson, 2009; Gailmard and Patty, 2012) and empirically (see Vogler (2022); Dahlström and Lapuente (2022) for two reviews). We contribute to this literature by estimating a production function of legislative output involving both politicians and a team of appointees.

Section 2 of this paper provides additional background on personal staffers in the US House and the legislative process. Section 3 lays out the datasets used. Section 4 contains a framework for estimation in team settings where individual contributions are unobserved, and is self-contained for readers interested in applying this framework.[6] Section 5 estimates the staffer share of differences using a linear decomposition, a dynamic specification, and Bayesian estimation. Section 6 extends the Bayesian analysis to uncover the role that jobs, experience, diversity, within-team interactions, and ideology play in shaping staffers' legislative impact. Section 7 concludes.

2 Background

2.1 Personal staffers in the House

The modern US Congress employs over 17,000 staffers each year, spread across the many functions of Congress.[7] In this paper, we focus on the personal staff of Representatives

[6] Appendix B contains more estimation implementation details.

[7] These include personal staff to members of Congress in both the House and Senate, committee staff in both the House and Senate, agency staff for the Congressional Budget Office, Government Accountability

of the House. These personal staff perform a variety of functions for their office: they respond to constituent concerns by phone or (e-)mail, they write speeches and other public-facing communication, they meet with lobbyists and organize their Representative's schedule, and they write bills that may someday become law.

Staffer turnover is relatively high. The hours tend to be long, and pay relatively low. In Table 1, we show that the average staffer is compensated at just under $45,000 per year. Given the difficulties of the job, staffers often exit Congress after a few years of work. The average staffer in our sample stays in Congress for just over 2 years, and often leaves the Hill for lobbyist jobs. Given the high turnover in staffers, Representatives are constantly on the search for new staffers to hire and fill old roles. Priority in hiring is typically given to anybody with experience on the Hill, with private mailing lists distributing job opportunities first before the general public is notified of job openings.[8]

Offices in Congress tend to be highly hierarchical. We show a typical org chart for a Representative's office in Figure 1. Immediately reporting to the Representative themselves is the Chief of Staff, who in turn leads Directors in multiple domains—Legislation, Communication, Outreach, etc. The Legislative Director (LD) typically has Legislative Assistants or Legislative Aides (LAs) working for them, while Legislative Correspondents (LCs) perform the lowest level of work. Oftentimes, this is responding to constituent mail, although LCs may also conduct background research or assist with the drafting of legislation. Promotion up the job ladder often takes place within a year or two of beginning the job; in our data, the average staffer holds a job title for around 1.3 years.

Each Representative is allowed to hire a maximum of 18 full-time staffers on a fixed budget (known as a Member Representational Allowance).[9] Combined with the strict hierarchy of job titles, this can make promotion within an office difficult (if not impossible) if the staffer occupying the job above does not leave; thus, staffers often seek opportunities in other offices for promotion. Although ideology may play some role in these moving decisions, conditional on a Representative's political party, many staffers note that opportunities for upward mobility represent the most important reason for switching offices.[10]

Office, and Congressional Research Service, as well as capitol police, legislative clerks, leadership staff, and others. See Cross and Gluck (2020) for a review of bureaucrats in Congress.

[8]Even entry into a full-time role in Congress is often predicated on securing an internship at Congress first.

[9]In contrast, Senators may hire as many staffers as they like, although they are also subject to a budget constraint.

[10]This has been described to us in multiple interviews with former staffers.

2.2 The Legislative Process

We provide a very brief summary of the legislative process in the US Congress, focusing on bills that originate in the House.[11] Before a bill is introduced to Congress, it is typically written by legislative staffers from one or more offices.[12] Upon introducing the bill to Congress, the Representatives of these offices become known as the "original cosponsors" of the bill. The bill is then assigned to a committee, where it may be scheduled for markup (amendments), debates, and hearings, before eventually receiving a committee vote. Most bills die in committee. However, if a bill makes it through committee, it moves to the House floor where it is scheduled for debate and a roll call vote. A bill that passes the House roll call vote is sent to the Senate to go through a similar procedure.[13] If a bill passes both chambers of Congress, it is sent to the President, who may either sign the bill into law or veto it.

3 Data

3.1 Statements of Disbursement

The primary dataset that we rely on is the House's Statements of Disbursement (SoD), which record all spending conducted by the House. Critically, these include payroll records that we use to identify the staffers who work for each office at a given point in time. The SoD has been published electronically on a quarterly basis since 2016. We utilize complete data on staffers from 2016 to 2022.

From 2.9 million payments made by Congress, we extract a total of 732 representatives and roughly 15,000 staffers.[14] We identify roughly 2,800 staffers that switch offices ("movers") in our main sample and construct panel data, associating a staffer with a single office in each quarter. A detailed description of the data construction process is outlined in Appendix A. We also use this payment data to construct a few other measures of

[11] For a more detailed but digestible version of the process, see `https://www.congress.gov/legislative-process`. This process is explained in depth in Smith et al. (2013). Wilkerson et al. (2015) traces where policies come from across legislation.

[12] On occasion, bills are also given to staffers by lobbyists. These may be modified before introduction to Congress.

[13] If the bill was first introduced in the Senate and passed the vote there, then it will be introduced to the House with sponsors (but not original co-sponsors), and follow the procedure outlined above. Occasionally, bills are simultaneously introduced to both chambers of Congress, in which case a bill will still have an original co-sponsor in the House.

[14] We disambiguate names using LinkTransformer (Arora and Dell, 2023). This is manually validated in the case of all Representatives; accuracy is generally high (over 99%).

office spending, including usage of the franking privilege, spending on advertising, and spending on technology (including spending on computers, software, and technology service contracts).

We code the job title of each staffer in each quarter following the keyword coding procedure in Crosson et al. (2021). We classify jobs into five categories: legislative staff, political management, communications, office management, and constituency service. We additionally code several job titles that are of relevance to the legislative process: Chief of Staff, Legislative Director, Legislative Assistant/aide, and Legislative Correspondent.

We report summary statistics in Table 1. We find that roughly one-quarter of staffers work in a legislative role at any given point in time. In general, movers are not a representative sample of all staffers. They are disproportionately likely to have longer careers, higher wages, be in legislative positions, or to hold any sort of management title. Given that we are primarily interested in understanding the importance of staffers on the whole relative to members of Congress, rather than in the effectiveness of the average staffer, we view this as a feature rather than a bug. The fact that movers are likelier to hold important positions in an office makes us likelier to recover the true impact of staffers. However, the fact that we exclude a substantial fraction of staffers from the main analysis means that we are likely to underestimate the true effect of staffers.[15]

3.2 Other data sources

We supplement the House SoD with a variety of other data sources.

First, we collect bill activity from the Congress.gov API. We use this data to construct our primary measures of legislative productivity: the number of original co-sponsored bills by each Representative in each quarter, the number of these co-sponsored bills that exit committee, the number of these co-sponsored bills that pass a House floor vote, and the number of these co-sponsored bills that are signed into law.

We collect two sets of measures of legislative productivity from the Center on Effective Lawmaking. First, we make use of their measure of "important" bills, defined as bills that receive an end-of-the-year write-up in the Congressional Quarterly Almanac. Second, we use two additional measures of legislative effectiveness: (1) the Legislative Effectiveness Score (LES), a measure of the number of bills that an office produces, weighted by the importance of the bill and its legislative success, and (2) the LES minus a benchmark score that is a Representative's expected LES based on their seniority, chair status, and their party's majority status (Volden and Wiseman, 2009).[16]

[15] In our Bayesian analysis, we are able to bring the non-movers back into the analysis and confirm this.

[16] More details on both sets of measures are provided in Appendix A.

Finally, we collect data on roll call votes and DW-Nominate scores from voteview.com (Boche et al., 2018; Poole and Rosenthal, 1985). We also collect staffer-level covariates from LegiStorm, including gender, education, and work experience, and impute ethnicity from names in a procedure described in Appendix A.

4 A Mover Design for Team Environments

In this section, we present a framework for estimating the staffer share of differences between high and low productivity offices.

4.1 Linear decomposition

In the canonical mover design, an individual level outcome such as wages (Abowd et al., 1999), mortality (Finkelstein et al., 2021), or voting behavior (Cantoni and Pons, 2022) is estimated using a linear combination of individual fixed effects and group (often location) fixed effects:

$$y_{ijt} = \alpha_i + \gamma_j + \tau_t + \varepsilon_{ijt}$$

where i is an individual, j is a group, and t is a time period. In our setting, i is a staffer, j is a Representative's office, and t is a quarter.

Unfortunately, there are many circumstances in which group-level outcomes are influenced by individuals, but the specific contribution of each individual to the group is unknown, making this equation impossible to estimate. For instance, in our setting, we do not observe the specific legislation that each staffer produces. However, given that we observe the movements of staffers across different offices, as well as the legislative output of offices during these moves, we can exploit the information contained in these moves to inform our estimates of each staffer's productivity. Thus, we extend the standard mover design to team environments, continuing to assume a linear value-added model, or production function, of legislation.[17] Let $m_{i,j,t} = \mathbb{I}[\text{staffer } i \text{ in office } j \text{ at time } t]$. Then our model is:

$$y_{jt} = [\sum_i \alpha_i m_{i,j,t}] + \gamma_j + \tau_t + \varepsilon_{jt} \tag{1}$$

where m and y_{jt} are data, and the remaining terms are parameters to be estimated. We assume that $E[\varepsilon_{jt}|\{i\}_{m_{ijt}=1}, j, t] = 0$.

Our estimand of interest is the share of differences in legislative output between high and low productivity offices that is attributable to staffers. Under the standard AKM

[17] We relax the linearity assumption in Section 6 and estimate a nested CES model, among others.

mover design, this can be identified via a linear decomposition. A similar argument holds when we estimate the team mover design in Equation 1.

Let $\bar{y}_j = \gamma_j + \bar{y}_j^{\text{staff}}$ be the expectation of y_{jt} over all time periods, where

$$\bar{y}_j^{\text{staff}} \equiv \frac{1}{T} \sum_t \sum_i \alpha_i m_{i,j,t}$$

Hence, for any two offices d, o:

$$\bar{y}_d - \bar{y}_o = (\gamma_d - \gamma_o) + (\bar{y}_d^{\text{staff}} - \bar{y}_o^{\text{staff}}) \tag{2}$$

Equation 2 shows that the difference in average office productivity between two offices can be decomposed into two parts: the first part is the difference between the office fixed effects $\gamma_d - \gamma_o$; and the second part is the difference between staffer-specific components $\bar{y}_d^{\text{staff}} - \bar{y}_o^{\text{staff}}$. The share of the difference in office productivity between the origin and destination office that is driven by staffers and offices (Representatives) is therefore:

$$S^{\text{rep}}(d, o) = \frac{\gamma_d - \gamma_o}{\bar{y}_d - \bar{y}_o},$$

$$S^{\text{staff}}(d, o) = \frac{\bar{y}_d^{\text{staff}} - \bar{y}_o^{\text{staff}}}{\bar{y}_d - \bar{y}_o} = 1 - S^{\text{rep}}(d, o)$$

This allows us to estimate the staffer share of differences between any pair of offices. Throughout the paper, we estimate S^{rep} and subtract it from 1 for a consistent estimate of the staffer share S^{staff}.

We focus on estimating the difference between two sets of offices O and D—for instance, offices with above and below median legislative output. For an office level quantity x_d, let $x_D = \frac{1}{|D|} \sum_{d \in D} \bar{x}_d$. Then the Representative and staffer share of differences is simply:

$$S^{\text{rep}}(D, O) = \frac{\gamma_D - \gamma_O}{\bar{y}_D - \bar{y}_O} = 1 - S^{\text{staff}}(D, O) \tag{3}$$

4.2 Two dynamic estimating equations

Estimating staffer shares dynamically In order to track changes in office productivity around the timing of a staffer move, we can estimate an event-study version of Equation 1. To do so, we extend the derivation in Cantoni and Pons (2022) to a team setting. Let $r(i,t) = t - t_i^*$ be the time to a staffer i's first move and $\rho_{r(i,t)}$ be the time fixed effect relative to the move. Then we can modify Equation 1 as follows:

$$y_{jt} = \alpha_i m_{ijt} + [\sum_{i' \neq i} \alpha_{i'} m_{i'jt}] + \gamma_j + \rho_{r(i,t)} + \tau_t + \epsilon_{jt} \tag{4}$$

Consider a staffer i who moves from an origin office $o(i)$ to a destination office $d(i)$.

Let $\delta_i \equiv \bar{y}_{d(i)} - \bar{y}_{o(i)}$, the difference in average office productivity between the destination and origin office of staffer i. Equation 4 can be rewritten as:

$$y_{jt} = \alpha_i m_{ijt} + \gamma_{o(i)} + \mathbb{I}_{r(i,t)\geq 0} \times S^{\text{rep}}(d(i), o(i)) \times \delta_i + [\sum_{i' \neq i} \alpha_{i'} m_{i'jt}] + \tau_t + \rho_{r(i,t)} + \epsilon_{ijt}$$

Let $\tilde{\alpha}_i = \alpha_i + \gamma_{o(i)}$, substitute $\mathbb{I}_{r(i,t)\geq 0}$ with indicators for time to move, and substitute δ_i with its sample analogue $\hat{\delta}_i = \hat{\bar{y}}_{d(i)} - \hat{\bar{y}}_{o(i)}$. The event-study specification becomes:

$$y_{jt} = \tilde{\alpha}_i m_{ijt} + \theta_{r(i,t)} \hat{\delta}_i + [\sum_{i' \neq i} \alpha_{i'} m_{i'jt}] + \tau_t + \rho_{r(i.t)} + \epsilon_{jt}$$

where $\theta_{r(i,t)}$ is the Representative share of differences between offices at $r(i,t)$. This specification is computationally expensive to estimate.[18] This motivates the use of the following outcome:

$$\hat{y}_{it} = \mathbb{I}_{r(i,t)\geq 0}[y_{d(i)t} - \sum_{i' \neq i} \hat{\alpha}_{i'} m_{i',d(i),t}] + \mathbb{I}_{r(i,t)<0}[y_{o(i)t} - \sum_{i' \neq i} \hat{\alpha}_{i'} m_{i',o(i),t}]$$

and event study specification:

$$\hat{y}_{it} = \tilde{\alpha}_i + \theta_{r(i,t)} \hat{\delta}_i + \tau_t + \rho_{r(i.t)} + \epsilon_{it} \tag{5}$$

Notice that Equation 5 is equivalent to the prior event study specification, provided that $\alpha_i = \hat{\alpha}_i$.[19] It has the advantage of being estimable by standard fixed effect routines as we no longer have a variable number of fixed effects per observation to estimate. However, it requires an unbiased estimate of $\alpha_i, \hat{\alpha}_i$. Reusing $\hat{\alpha}_i$ from the same data is problematic as it uses the data twice to validate the estimate. Instead, we split our sample into odd quarters and even quarters. We first estimate Equation 1 above on odd quarters only. Then, we turn to even quarters using the fixed effects estimates from odd quarters to estimate Equation 5.

This delivers the estimand of interest $\theta_{r(i,t)}$, the share of differences between the destination and origin offices that are attributable to the Representative.

Predicting a staffer's movements We continue to analyze a staffer's first move between origin office o and destination office d. We develop a new estimation strategy to validate

[18] As discussed in Appendix B, we estimate Equation 1 using the full design matrix containing indicators for every staffer and $O(JT)$ rows of data. However, once we move to an event study specification, we have $O(IT)$ rows of data and $I >> J$.

[19] In the case where $\alpha_i = \hat{\alpha}_i$, the errors $\epsilon_{it} = \mathbb{I}_{r(i,t)\geq 0} \epsilon_{d(i),t} + \mathbb{I}_{r(i,t)<0} \epsilon_{o(i),t}$. This motivates clustering errors at the representative level.

Equation 1 in a dynamic fashion. Specifically, we estimate:

$$y_{d(i),t} = \sum_{\tau=-4}^{4} \mathbb{I}[r(i,t-\tau)=0] \left(\theta_\tau \hat{\alpha}_i + \eta_\tau \left(\sum_{i' \neq i} \hat{\alpha}_{i'} m_{i',d(i),t} \right) + \zeta_\tau \hat{\gamma}_{d(i)} \right) + \rho_{r(i,t-\tau)} + \hat{\tau}_t + \varepsilon_{d(i),t} \tag{6}$$

where $\theta_\tau \hat{\alpha}_i$ captures the specific contribution of a staffer i to the destination office's output at each time lag τ, $\eta_\tau \left(\sum_{i' \neq i} \hat{\alpha}_{i'} m_{i',d(i),t} \right)$ captures the aggregate effect of other staffers on the destination office's output at each time lag τ, and $\zeta_\tau \hat{\gamma}_{d(i)}$ captures the time-varying Representative-level fixed effect at each time lag τ.[20]

If the model in Equation 1 is true, then we should expect the following to hold: (1) θ_τ should be equal to 0 prior to the move (because the staffer is not in the destination office), and equal to 1 after the move (as they contribute exactly their fixed effect after moving), (2) η_τ should always equal 1, because the team of staffers at the destination office is always present at the destination, and (3) ζ_τ should always equal 1, because the destination office Representative is always in the destination office. Note that if these conditions hold, then Equation 6 directly collapses to Equation 1, except stacked at the individual level.[21]

Estimating this equation serves as a validation of the model in three ways: (1) a $\theta_\tau > 0$ means that we can predict when a staffer has moved based purely on whether their fixed effect appears to be contributing to the office's output. Since $\hat{\alpha}_i$ is estimated out of sample, this signals that an individual fixed effect is estimated with relative precision, (2) an η_τ and ζ_τ that remains constant over time serves as both a placebo test and a validation that these fixed effects are well estimated, and (3) pre-trends can be tested across θ_τ, η_τ, and ζ_τ for $\tau < 0$. Thus, we can test for anticipatory behavior in the staffer, their team, and the Representative all at once. Finally, we note that by replacing the destination office with the origin office in Equation 6, one can examine whether staffers have a persistent effect on offices even after they leave.

4.3 Bayesian estimation

We next move to estimating the staffer share of differences using hierarchical Bayes. Specifically, we re-estimate Equation 1, assuming that $\{\alpha_i\}_i, \{\gamma_j\}_j$ are drawn from a common distribution of staffer and Representative productivities respectively. We estimate this model using Hamiltonian Monte Carlo (HMC), a type of Markov Chain Monte Carlo technique. This method makes the estimation of problems with difficult geometries and

[20] We face the same problem as the prior dynamic estimation if the same data is used to estimate fixed effects α_i, γ_j as well as the main event study. Thus, we estimate these fixed effects on odd quarters and estimate Equation 6 in even quarters.

[21] This once again motivates clustering standard errors at the Representative level.

large numbers of parameters tractable.[22] With estimates of these random effects in hand, we can follow the procedure outlined in the sections above to derive the staffer share of differences.[23]

We adopt a Bayesian approach for a few key reasons. First, in frequentist settings, estimating a large number of parameters leads to issues with consistency. Our data-generating process is such that the number of parameters (individual fixed effects) grows with the size of the dataset, but we do not gain more information about individuals as the dataset grows—collecting more years of data does not help when an individual has already left Congress. As the number of parameters increases, the errors in these estimates do not necessarily average out, leading to biased or inconsistent parameter estimates (Andrews et al., 2008). On the other hand, Bayesian estimation provides more stable and consistent parameter estimates by allowing for estimation of the distribution of parameters that α_i and γ_j are drawn from. This overcomes the consistency problem discussed above, as well as issues like overfitting and multicollinearity that can arise when the number of parameters to be estimated is large relative to the sample size.[24]

Second, Bayesian estimation offers more flexibility than OLS, allowing us to explore a richer space of nonlinear models. With Bayesian models, we can incorporate job information to evaluate how staffers' job titles influence their legislative productivity, estimate how experience in Congress impacts productivity, incorporate CES functions to analyze how staffers and Representatives might substitute or complement one other, and study how staffers and Representatives both affect the ideological content of bills.[25]

Third, we depart from the empirical Bayes approach used by many economists in favor of a pure Bayesian approach.[26] This is because empirical Bayes is dominated by pure Bayes, in a sense made precise by a number of Complete Class Theorems.[27]

4.4 Discussion

In a recent review of personnel economics, Hoffman and Stanton (2024) note that while "understanding teams and team production is of huge practical importance," the actual

[22] We describe HMC in more detail in Appendix B. To the best of our knowledge, the only other paper in economics that makes use of HMC is Childers et al. (2022).

[23] We leave a discussion of specific distributional assumptions and the prior to Section 5.

[24] For more work on correlated random effects in the mover design, see for instance Bonhomme et al. (2019).

[25] We discuss specific model extensions in more detail in Section 6.

[26] For some examples of empirical Bayes used in value-added models, see examples such as Chetty et al. (2014); Angrist et al. (2017); Chetty and Hendren (2018). Purely Bayesian approaches have also been used before, as in Abowd et al. (2015).

[27] For more details, see Ferguson (1967); Lehmann et al. (1986); Berger (2013).

literature on teams is "relatively sparse." We compare our approach to other leading proposals in the literature, focusing on performance in our context, where teams may contain up to 18 staffers.[28]

Chan (2021) uses an empirical Bayes procedure to estimate doctor productivity in a setting where junior and senior physicians are paired together in teams of two. Extending this estimation framework to our setting requires the estimation of between 1.2 million and 1 quadrillion parameters, which is infeasible in our context where we lack this many observations.[29] Bergeron et al. (2022) and Bonhomme (2021) both use a linear model to estimate individual effects. Their linear models are similar to ours and estimated similarly; our contribution is to focus on a mean decomposition, which also allows us to estimate two novel dynamic specifications. Bonhomme (2021) also estimates a nonlinear model via a variational Bayes (ELBO) procedure, restricting attention to teams no larger than two.[30] The ELBO procedure in our setting is infeasible largely for reasons of computational efficiency: a single call to the objective function requires over 1 trillion basic operations. Finally, Ahmadpoor and Jones (2019) provide an estimation routine that uses simple gradient descent to solve a nonlinear least squares problem: recovering individual effects in a CES production function. While their method is computationally feasible (albeit around 100 times slower than ours), it is numerically quite unstable. For instance, even in an ideal case where we initialize all parameters to the CES problem with the oracle solution (with $\rho_{\text{staff}} = 0.794$), the method returns a result that is substantially different ($\hat{\rho}_{\text{staff}} = 9.461$). On the other hand, in Appendix B, we validate our procedure via simulation, showing that our dynamic and Bayesian estimates do not suffer from the same problems.

In addition to the approaches above, researchers have also adopted misspecified models that fail to correctly account for the impact of teams. This typically involves conducting analysis at the individual level or only focusing on one key individual per team.[31] In Section 5, we discuss how such misspecification would change our findings.

Thus, although we focus on estimating the effect of Congressional staffers in this paper, we hope that the framework laid out in this section will be useful for many settings. In domains ranging from healthcare to innovation to finance to the workforce at large,

[28] In our entire sample, the average team contains 12.94 staffers. In the mover sample, the average team contains 4.91 staffers.

[29] The specific number of parameters depends on how one extends the model. More details on this and the other papers discussed in this paragraph are described in Appendix B.

[30] Bonhomme (2021) writes that "it will be important to devise computational strategies to efficiently handle larger teams."

[31] Constantine and Correia (2021) list many examples of papers that are misspecified in this way. They provide an implementation of the linear model in Equation 1, but because their procedure cannot recover the FEs (individual FEs are treated as nuisance parameters), it cannot be used for decompositions.

there are many team environments where individual contributions are unobserved and this framework may be applicable.

5 The Staffer Share of Legislative Productivity

How important are staffers relative to members of Congress? We answer this question by estimating the share of differences in legislative productivity between high and low productivity offices that is attributable to staffers. We adopt three sets of approaches to estimating this quantity: we first estimate a series of static linear decompositions, then estimate the staffer share dynamically, and finally use a Bayesian approach to estimate the staffer share.

5.1 Linear decompositions

We begin by estimating a series of linear decompositions following Equation 1 of our mover sample. This allows us to estimate staffer effects (α_i) and Representative effects (γ_j) for each staffer and Representative in the sample. In each quarter, we identify the above and below median offices for five legislative outcomes of interest: the number of original co-sponsored bills introduced by the office, the number of these bills that exited committee, the number of these bills that passed a House floor vote, the number of these bills that were signed into law, and the number of "important" bills introduced.

We next calculate the representative share of differences between these high and low productivity offices, S^{rep}, following Equation 3, and subtract it from 1 to obtain a consistent estimate of the staffer share of differences, S^{staff}. These staffer shares are plotted in the blue bars of Figure 2. We form confidence intervals using the Bayesian bootstrap.[32]

We find that staffers are a critical input to the most basic output of Congress: bills and laws. 48.9% of the difference in bills produced between above and below median offices can be attributed to staffers. As we move across the life-cycle of a bill, we find a U-shaped pattern for the importance of staffers. Staffers account for only 20 to 25% of the difference in the number of co-sponsored bills that exit committee or pass a House floor vote. Given that each committee has its own staff who handle the text of bills, and the importance of political considerations in ushering a bill out of committee and through a House vote, we find the smaller role of personal staffers in these intermediate stages of a bill's life-cycle to be plausible. However, the importance of staffers appears to tick up again, to roughly 40%, once we examine the share of differences in co-sponsored bills that

[32]More details on the exact estimation procedure are contained in Appendix B.1.

eventually become law. After the passage of a bill through the House, it must also pass the Senate and be signed by the President. Inter-branch negotiation and coordination may require more manpower and connections, both of which personal staff may be able to provide.

We also examine the staffer share of differences in the number of important bills co-sponsored by the office. We find that staffers play a much smaller role in the introduction of these important bills, with a point estimate of 14.1% and a staffer share of 0 falling within the 95% confidence interval. We attribute this to several potential factors: first, it is likely that an office's most substantive bills are receive disproportionate attention from the Representative, who may otherwise trust their staff to handle regular bills. Second, bills are classified as "important" if they receive a year-end write-up in the Congressional Quarterly Almanac; a bill may be likelier to receive such a write-up if the Representative is spearheading the bill and promoting or advertising it, compared to bills originated from the staff. Finally, we note that bills classified as "important" are likely to be the subject of intense political dispute. An effective bureaucracy is likely to focus on issues of policy rather than politics (Weber, 1968), which may make the bills less contentious and worthy of a write-up.

Robustness, IV, and discussion For robustness, we also compute the staffer share of the difference between top tercile and bottom tercile offices (plotted in green) and between top quartile and bottom quartile (plotted in purple). We find that our estimates are very similar, even when looking at the more extreme differences between offices.

We perform two validations of the staffer fixed effect estimates. First, we regress the absolute value of each staffer's fixed effect $|\alpha_i|$, to capture the impact of a staffer (good or bad), against their wage w_i in Table 2, Panel A. We find that a standard deviation increase in a staffer's impact is associated with a 5 to 10% standard deviation increase in wages, consistent across outcomes. An attentive reader may notice that this regression coefficient is not identified: the fixed effects of two staffers with identical moving histories cannot be separately identified.[33] We take several steps to address this. In Panel B, we restrict attention to only staffers with unique work histories (over 95% of the sample) and find extremely similar results. In Panel C, we collapse the data to the work history level (taking an average across wages) and find similar results. Weighting observations by the number of underlying staffers in Panel D also does not materially affect the estimate. To demonstrate that this correlation is not driven by the difference between major job titles,

[33]While this regression coefficient is not identified, we emphasize that the staffer share of differences is, due to the linearity of the underlying decomposition. Once we move to non-linear models, we use Bayesian estimation, where the distribution on α_i disciplines the staffer random effect estimates.

in Panels E and F, we focus on the subset of staffers that are managers, or managers and legislative staff. We find that, if anything, the relationship between a staffer's impact and their wage is stronger within these samples.

Second, we correlate staffer fixed effects across specifications and present results in Table A.1. Despite the variation in staffer shares across outcomes, it appears that productive staffers are helpful across all stages of the bill life-cycle: the within-staffer correlations are consistently larger than 0.50 and statistically significant.

We also test for the staffer share of legislative productivity using the Center for Effective Lawmaking's Legislative Effectiveness Score (LES) and the LES minus a benchmark of Representative characteristics (including seniority, chair status in committee, and party in the majority). We present results in Figure A.1. We estimate that the staffer share of differences in legislative effectiveness is roughly 40% in both outcomes, consistent with our findings of the staffer share in bill production and passing laws.

We stress the importance of using the team-based design in Equation 1 by comparing it to misspecified approaches that fail to properly account for the presence of teams. In Table A.2, Panel A, we present our baseline estimates from the team-based mover design. Panel B presents estimates from the standard AKM design at the individual level. Across outcomes, the estimated staffer share is at most half as large as the baseline. Focusing on key individuals within the team does not help matters: in Panels C and D, we restrict attention to Chiefs of Staff and Legislative Directors respectively, and find that the staffer share of differences is similarly underestimated.

One may also be concerned about the exogeneity of the mover design. Although qualitative work suggests that opportunities up the job ladder represent the dominant reason for moving between offices, there are many potential reasons as to why staffer productivity may be correlated with their movement.[34] We test for persistent or slow-moving sources of endogeneity in Section 5.2 and do not find evidence of endogeneity. This leaves the possibility of fast-moving sources of endogeneity. We address these using an instrumental variables strategy.

We deploy two sets of instruments for staffer movements between offices: first, we exploit the fact that following turnovers in Congress, staffers often move to other offices within the same state (a home-bias effect), and use an indicator for a staffer's origin office losing an election and the potential destination office being in the same state as an instrument. Second, we draw on the migration literature and interact a "push" factor be-

[34]For example, more productive staffers may be poached by more productive offices, while less productive staffers may be disproportionately fired by some offices; staffers may have ideological preferences in choosing offices to work for, which may be correlated with productivity; staffers may go to offices where they think they will be more impactful.

tween two offices (the leave-out total flow of staffers between the origin and destination) with a "pull" factor into the destination (the number of staffers who leave the destination in that period). Given the high number of moves that we are instrumenting for, a very strong first stage is required. To achieve this, we deploy machine learning methods to estimate the first stage.[35] We present results from this estimation procedure in Figure A.2. The IV estimates deliver more extreme staffer shares: the staffer share in co-sponsored bills hovers around 80%, while the later stages of the bill life-cycle have point estimates much closer to 0%. Interestingly, the staffer share of important bills also rises to 60-80%. It is worth noting that the standard errors increase substantially, and for certain stages of the bill life-cycle (bills exiting committee or passing a House vote), the confidence interval covers virtually all plausible values.[36] Nonetheless, we are unable to reject the hypothesis that the OLS and IV estimates are different.

Finally, we discuss the interpretation of the estimated staffer shares. While we term $S^{\text{staff}}, S^{\text{rep}}$ the staffer share and Representative share respectively, in reality, they capture all forces that are always associated with the staffer or the Representative. Thus, if there are non-mover staffers who are always attached to the same Representative throughout our sample, we will attribute this staffer's impact to the Representative share, and hence likely underestimate the true staffer share. Similarly, a Representative may appear effective not because of their inherent quality, but because of external factors (for instance, they happen to be a committee chair for the entirety of our sample). However, Figure A.1 suggests that these factors are not quantitatively important in contributing to the Representative share of differences. Finally, lobbyists and other actors may follow staffers rather than Representatives, and provide them with policy ideas or literal bills over the course of their career; the impact of these people will be bundled into the estimated staffer fixed effects. As such, the staffer share that we compute should be interpreted as representing the staffers' contribution broadly construed, inclusive of all connections or other permanent factors associated with them.

5.2 Dynamic estimation

We proceed by estimating two sets of dynamic specifications. This allows us to map the pathway of staffer influence over time, as well as test for potential pre-trends and validate our linear decomposition estimates.

[35] The complete estimation procedure is described in Appendix B.1.

[36] Given the lack of precision in these estimates, we prefer to interpret the IV estimates as a sanity check that may help us to sign bias, rather than relying on the specific quantitative estimate.

Dynamically estimating the Representative share of differences in legislative productivity We estimate event studies of the Representative's share of differences in legislative productivity between a staffer's origin and destination office. Throughout our dynamic analysis, we focus on the first move that a staffer makes to simplify the interpretation of the estimated quantities. As noted in Section 4, we estimate all fixed effects on odd quarters and estimate the event study Equation 5 on even quarters. In Figure 3, we plot the estimated $\theta_{r(i,t)}$ coefficients, along with 95% confidence intervals with underlying standard errors clustered at the Representative level.

We find no evidence to support pre-trends: across all specifications, the estimated coefficient is never different from 0 with statistical significance at any time period. The difference between a staffer's origin and destination office productivities $\hat{\delta}_i$ are not correlated with the origin office's productivity $\hat{y}_{ijt}$ (which includes the staffer's own productivity α_i). This supports the causal interpretation that we advance, as it rules out slow-moving sources of endogeneity (such as a staffer becoming more productive over time, being noticed by a high-productivity office, and hired for that reason).

We note that there is a sharp increase in the Representative share in the exact quarter that the staffer moves in all specifications. This share then stabilizes for the remaining quarters. In Figure 3, Panel A, we find that the estimated Representative share of differences in bills produced is typically around 60-70%, although the confidence intervals frequently include the 50% quantity estimated in the linear decomposition.[37] In Panels B-D, we find that the Representative share of differences through the bill life-cycle also hovers around the 50-60% range, implying a staffer share of 40-50%. In Panel E, examining important bills, we once again find that the Representative share is close to 100% (and the staffer share close to 0%).

The relative stability of coefficients across quarters suggests that staffers immediately have an impact upon moving offices—it does not take multiple quarters for their presence to be felt. We additionally validate this finding with an in-sample estimation on the odd quarters. We present results in Figure A.3, noting the absence of pre-trends, the immediate jump in Representative shares in the quarter the staffer moves, and the stability of the coefficient in the periods afterward.

Taking these results on the whole, we once again find that staffers play a substantial role in the legislative process.

[37] We note that the underlying shares estimated are different quantities: in the linear decomposition, this is the share of differences between high and low-productivity offices, while in the dynamic specification, this is the share of differences between destination and origin offices. Staffers do not systematically move to higher productivity offices.

Dynamically estimating staffer movement Instead of estimating the Representative share of differences in legislative productivity, we can also dynamically predict when staffers move offices. We take the estimated Representative and staffer fixed effects from odd quarters and estimate Equation 6 in even quarters. In Figure 4, we plot the estimated $\theta_\tau, \eta_\tau, \zeta_\tau$ coefficients, along with 95% confidence intervals with underlying standard errors clustered at the Representative level.

We first note the absence of pre-trends across outcomes in a staffer's own fixed effect (θ_τ, plotted in red), once again supporting a causal interpretation. Prior to a staffer moving, there is no correlation between their own productivity α_i and the destination office's output $y_{d(i),t}$. Strikingly, we are able to predict the presence of a staffer moving into an office using only their (out of sample) fixed effect. Immediately upon moving, we are able to detect the presence of the staffer from their impact on legislative productivity, and in Panels A-C (where the outcomes are co-sponsored bills, bills exiting committee, and bills passing a House vote), we are able to reject the null of the mover having no impact on the office. This effect is persistent and stable over time.[38] Unfortunately, we are unable to reject the null in Panels D and E (bills becoming law and important bills), largely because these are low-frequency events that we are underpowered to detect.[39]

We next turn our attention to the office fixed effect (ζ_τ, plotted in blue) and the other staffers' fixed effect (η_τ, plotted in grey). Following the theoretical prediction of the model, we find that these coefficients remain relatively stable over time, exhibiting only minor deviations from the normalized coefficient at $\tau = -1$ in both the pre- and post-period. This also lends empirical support to the model that we estimate.

One may wonder why the estimated staffer's own effect θ_τ differs from 1 when a literal interpretation of the model would suggest that it should be precisely 1 in the post-period. This may be for several reasons: first and foremost, there may be substantial measurement error in the estimated fixed effects, which should attenuate any of our coefficient estimates. Second, our model may be misspecified—assuming linearity across all staffers and the Representative is a strong assumption, and one that we explicitly test in Section 6. We analyze these possibilities by presenting the in-sample validation on odd quarters in Figure A.4. In-sample, we find that the own staffer coefficient θ_τ is now much closer to 1 in the post-period, indicating that measurement error may be substantial.

Finally, we estimate whether staffers have a persistent effect on their offices after they leave. To do so, we re-estimate Equation 6, except we now use the output from the origin

[38] We are aware that the third quarter after a move is consistently slightly different from other quarters, even with the inclusion of event-time fixed effects. We have not been able to find an explanation for this.

[39] We note that in Panel E, a joint test of the post-period own staffer fixed effects being different from 0 has a p-value of 0.103.

office $y_{o(i),t}$ as the outcome instead. We plot results out-of-sample and in-sample in Figures A.5-A.6. We find that staffers have no persistent effect on their offices: in the very quarter that they leave, we cannot distinguish their effect from 0, whereas the effect of the Representative and other staffers remains constant. Thus, the impact of staffers on legislation appears to be contingent on their physical presence, rather than any deeper structural or cultural changes they may impart on an office.

5.3 Bayesian estimation

We next estimate our original team-based mover design model, specified in Equation 1, using Bayesian methods. We employ fully Bayesian methods in estimation for several key reasons: First, we cannot guarantee consistency of fixed effects with frequentist inference. However, we can consistently and efficiently estimate the distribution of these effects under a hierarchical model with correlated random effects. Second, the Bayesian approach allows us to estimate more flexible models, a feature that we exploit in Section 6. Third, the empirical Bayes approach that is more popularly employed in the economics literature is dominated by fully Bayesian ones.

In our baseline specification, we assume that α_i, γ_j are distributed log-normally according to $(\mu_\alpha, \sigma_\alpha)$ and $(\mu_\gamma, \sigma_\gamma)$ respectively. As with all Bayesian estimations, the choice of priors is important. We select weakly informative priors that are designed only to rule out extremely implausible parameter estimates, and rely on the data to inform the posterior. For example, the number of bills co-sponsored by an office in a quarter has a mean of 16.3 and a standard deviation of 18.9. In our baseline specification, the prior for both μ_γ and μ_α is N(0,10), meaning that the average Representative or staffer producing 0.0001 bills or 10,000 bills are all within 1 standard deviation of the prior mean. We believe that this captures all reasonable levels of potential Representative and staffer productivity. In Appendix A, we provide a full specification of the priors employed and other estimation details. Throughout the remainder of this paper, we focus our attention on the number of co-sponsored bills produced by an office, given that we are most powered to detect effects for this outcome of interest.

We present a trace plot for the Representative share of the difference in co-sponsored bills between above and below median offices in Figure 5, Panel A. We also report the posterior mean, 2.5th, and 97.5th percentiles for the Representative share in Table 3, Panel A. We find that the Representative share of the difference is just over 60%, or that the staffer share is just under 40%, a very similar quantity to that estimated in the dynamic

specification (Figure 3, Panel A).[40]

In Figure 5, Panels B-D, and Table 3, Panel A, we present trace plots and summary statistics for the posterior for the distribution of staffer effects and Representative effects. All signs indicate model convergence. Specifically, we find posterior means of $(\mu_\alpha, \sigma_\alpha) = (-5.38, 3.38)$ and $(\mu_\gamma, \sigma_\gamma) = (2.52, 0.52)$. This implies that the average staffer produces $\exp(\mu_\alpha + \frac{\sigma_\alpha^2}{2}) = 1.39$ bills per quarter, and that the average Representative produces 14.22 bills per quarter. However, the median staffer produces roughly 0.004 bills per quarter, in line with the fact that the majority of staffers do not work in roles directly tied to legislative work. In Figure A.7, we plot the posterior distributions for $\log(\alpha_i)$ and $\log(\gamma_j)$. As expected, the vast majority of staffers have virtually no impact on the legislative process. However, there is a large right tail of staffer productivity, including a handful of staffers who have productivities on par with Representatives. On the other hand, the distribution of Representative productivity is much more concentrated, suggesting that there is much less dispersion in Representative legislative productivity.

Robustness In Figure A.8, Panel A, we replicate the trace plot for the Representative share of differences in co-sponsored bills for above versus below median productivity offices. In Panel B, we show the Representative share for top versus bottom tercile offices, and in Panel C, we show the Representative share for top versus bottom quartile offices. Examining these alternate differences between offices does not materially change the Representative share.

In Figure A.9, Panel A, we replicate the trace plot for the Representative share of differences using our baseline log-normal distribution. In Panels B-D, we show the Representative share using normal, Fréchet, and a mixture of normal distributions instead. All display similar (if not slightly larger) staffer shares to the baseline log-normal distribution.[41]

In Figure A.10, Panel A, we replicate the trace plot for the Representative share of differences in co-sponsored bills. In Panels B-E, we show the Representative share for our alternate measures of legislative productivity: co-sponsored bills exiting committee, co-sponsored bills passing a House vote, co-sponsored bills becoming law, and co-sponsored

[40] In Table 3, we also present $\tilde{n}$, the effective sample size of the estimated parameter, and $\hat{R}$, a measure of MCMC convergence (where closer to 1 is better). The effective sample size varies across parameters because it is decreasing in the within-chain correlation of sampled posterior values, which may vary by parameter. For $\hat{R}$, a common rule of thumb is that values beneath 1.05 are acceptable (see `https://mc-stan.org/rstan/reference/Rhat.html`). We present additional properties of the sampler in Table A.5. All diagnostics look normal.

[41] The mixture of normals captures the (potentially) bimodal nature of staffer productivities, between those who are involved in legislative work and those who are not. We run the mixture of normals distribution MCMC for longer to better assess model convergence.

important bills. We find that the Bayesian estimation tends to produce slightly larger staffer shares than OLS.[42]

While a variance decomposition cannot be conducted on the fixed effects estimated via OLS, they can be on the Bayesian posterior. In Table 4, we decompose the variance at the office level into a staffer component and a Representative component. We find that staffers are responsible for 37% of the variance, a figure slightly below our mean decompositions.[43]

Finally, a concern over the entire analysis presented thus far is that it does not account for staffers who never move. For instance, if non-movers have productivities that are negatively correlated with the movers, then it is possible that the staffer share is upwardly biased. We are able to account for these non-movers in our Bayesian analysis, simply by including the non-movers as well when estimating Equation 1.[44] In Figure A.11, we replicate Figure 5, now including all non-movers into the estimation as well. Panel A reveals that, as expected, accounting for this missing mass of staffers in the analysis pushes up the staffer share to roughly 65%.[45]

Taking stock of the linear decomposition, dynamic estimation, and Bayesian estimation, we find a consistent throughline of staffers accounting for 40% or more of the difference in co-sponsored bills produced between offices. Although estimates for other stages of the bill life-cycle and important bills are often less precise, we are typically able to bound the staffer share away from 0 and often find staffer shares comparable to that seen in the number of co-sponsored bills. Collectively, these results indicate that staffers play a substantial role in the production of legislation.

6 Unpacking the role of staffers in Congress

In this section, we examine the role that various factors may play in determining the impact of staffers.

[42]The exception is in Panel E, important bills, where the estimated staffer share is close to 80%—we note that this is similar to the IV point estimate, and still within the 95% confidence intervals of the dynamic estimates.

[43]The staffer and Representative variances sum to more than 100% due to the correlation between the two.

[44]Note that the distribution of effects of staffers who do not move, but do enter or exit the office (i.e. staffers who join or leave Congress during our sample), is identified. For intuition, consider a staffer who leaves an office at time t. If we see that the number of bills produced by the office substantially drops from $t-1$ to t, that informs the productivity of the staffer. The most substantive assumption underlying this analysis is that staffers who never move, enter, or exit are drawn from a common distribution to the other staffers.

[45]We focus on the movers for the remainder of our analysis, as the much larger number of non-movers compared to movers makes including them computationally expensive.

6.1 Jobs

Which jobs along the career ladder best facilitate the production of bills? We estimate an extension of the model in Equation 1 that allows for jobs to flexibly scale up (or down) the productivity of staffers based on their individual productivity. Specifically, we estimate the model:

$$y_{jt} = \sum_i \alpha_i \beta_{J(i,t)} m_{ijt} + \gamma_j + \tau_t + \varepsilon_{jt} \tag{7}$$

where $J(i,t)$ refers to the job that a staffer i holds at time t, and β is the new parameter of interest. We separate jobs into four categories based on the legislative staffer career ladder: Legislative Correspondent (LC), Legislative Assistant/Aide (LA), Legislative Director (LD), and Chiefs of Staff (CoS), as well as an "other" category for all other jobs. We normalize the β for all other jobs to 1.

We present the estimated β parameters in Figure 6. We find that most legislative jobs have a roughly similar amount of productivity (at 3-4 times the typical job), but that there is a sharp drop-off when a staffer is promoted to being a Chief of Staff in their office (down to 0.3). One reason for this may be that Chiefs of Staff are responsible for many different domains, ranging from legislation to communication to regular office administration, and may not be able to devote as much attention to legislative issues. As such, despite producing legislation being a top priority for many Congressional offices, it appears that the Chief of Staff has much less direct influence over the legislative output of an office when compared to more junior positions.[46] Indeed, the fact that the Chief of Staff coefficient is well below 1 suggests that even holding non-legislative jobs will generally allow a staffer to have greater influence over an office's legislative output than being the Chief of Staff!

6.2 Experience

We next examine the role that experience plays in determining an individual's legislative productivity. We estimate an extension of the model in Equation 1 that allows for jobs to flexibly scale up (or down) the productivity of staffers based on their individual productivity. Specifically, we estimate the model:

$$y_{jt} = \sum_i [\alpha_i m_{i,j,t} \tau_{total}(i,t)^{\phi_{staff}}] + \gamma_j (\tau_{total}(j,t)^{\phi_{rep}}) + \tau_t + \varepsilon_{jt} \tag{8}$$

[46]The coefficient for LCs is imprecisely estimated. This is likely due to the fact that LCs may be given different responsibilities in different offices. Some LCs only handle constituent mail, while others are involved in crafting legislation. Unfortunately, we cannot distinguish different types of LCs from the job title information that we have.

where τ_{total} reflects the total amount of time (in quarters) that a staffer or Representative has spent in Congress, and the exponent ϕ is the parameter of interest.

We present trace plots of the estimated ϕ_{staff} and ϕ_{rep} in Figure 7. We find that the posterior mean for ϕ_{rep} is 0.17, implying that Representatives with 1 full session of experience in Congress (8 quarters) are 42% more effective than a freshman with no experience, and that Representatives with 10 full sessions of experience in Congress (80 quarters) are 210% more effective than a freshman with no experience. On the other hand, we are unable to reject the null of staffer experience being irrelevant, $\phi_{\text{staff}} = 0$. We note that although our estimates are imprecise (a 95% confidence interval would just barely exclude gains comparable to Representatives, $\phi_{\text{staff}} > 0.17$), the lack of a detectable effect is consistent with Crosson et al. (2020), who find that total legislative staff experience does not correlate with lawmaking effectiveness.[47]

6.3 Diversity

We also examine the impact of staffer diversity on legislative productivity. To measure diversity, we calculate the share of staffers belonging to each major ethnic group (13 categories, such as "Greater African, Muslim" or "Greater European, British"), gender, educational degree (bachelor's or below, MA, MPP, JD, or other advanced degree), and type of work experience prior to Congress (private sector, policy-facing, local government, or federal government) and compute the Herfindahl-Hirschman Index (HHI) for each office in every quarter along these four measures. We then re-estimate Equation 1, adding each one of these diversity measures in turn as an additional covariate.

We report coefficients and 95% credible intervals in Table 5. For each of these measures, we find that less concentration (more diversity) is associated with more productive offices. Effect sizes are small but statistically significant: moving from the 1st quartile to the 3rd quartile in ethnic diversity increases bills produced by 0.85% of a standard deviation, while the effect is 0.27% of a standard deviation for gender, 0.78% of a standard deviation for education, and 0.28% of a standard deviation for prior work experience. Thus, while we find that offices benefit from diversity, we also rule out extremely large gains from diversity.

[47] Specifically, Crosson et al. (2020) finds that the average effect of legislative staff experience is indistinguishable from 0, but that experienced staff positively impact legislative productivity for committee chairs and inexperienced members of Congress.

6.4 Substitutes and complements

The linear model that we have assumed in Equation 1 may not be sufficiently flexible to capture the full nuance of interactions between staffers and Representatives. For instance, how does skill aggregate within the team of staffers—is it better to have a single superstar staffer or two decent staffers? Is the team of staffers a substitute or complement to Representative skill? To answer these questions, we estimate an extension of the model in Equation 1 that includes a nested CES production function:

$$y_{jt} = \left[\left[\sum_i \alpha_i^{\rho_{\text{staff}}} m_{i,j,t} \right]^{\rho_{\text{rep}}/\rho_{\text{staff}}} + \gamma_j^{\rho_{\text{rep}}} \right]^{1/\rho_{\text{rep}}} + \tau_t + \varepsilon_{jt} \tag{9}$$

where ρ_{staff} and ρ_{rep} are the parameters of interest. Specifically, ρ_{staff} governs the degree of substitutability of staffers within a team, while ρ_{rep} governs the degree of substitutability between the team of staffers and the Representative.

We present trace plots of the estimated ρ_{staff}, ρ_{rep}, and Representative share of differences in Figure 8.[48] Recall that under perfect substitutes $\rho = 1$, and that under Cobb-Douglas $\rho = 0$. We find that staffers within a team are highly substitutable for one another. The posterior mean for $\rho_{\text{staff}} = 0.79$.[49] This suggests that maximizing the sum of staffer productivities α_i is close to optimal, regardless of whether the skill is highly concentrated within one individual or spread over many. On the other hand, we find that the production function for the team of staffers and the Representative is close to Cobb-Douglas. The posterior mean for $\rho_{\text{rep}} = 0.13$.[50] Heuristically speaking, this means that an office consisting of a superstar Representative but very unproductive staffers and an office consisting of a very unproductive Representative and superstar staffers will both be less productive than an office containing a more balanced mix of skill between Representatives and staffers. Thus, Representatives facing a hiring trade-off between staffer compensation and staffer skill should hire staffers that tend to reflect their own productivity.

Identifying Pareto improvements Do Representatives generally follow this guideline? We formalize this by testing for potential Pareto improvements, circumstances under which two offices may trade a staffer and both are left better off after the trade. We operationalize this by additionally assuming that a Representative has access to a technol-

[48] We omit the Representative share of differences for the job and experience model extensions because they are extremely similar to the baseline estimate. Specific estimates are available on request.

[49] This translates to $\sigma = \frac{1}{1-\rho} = 4.76$.

[50] This translates to $\sigma = 1.15$.

ogy that produces bills with a linear cost function. One interpretation of this technology could be Congressional interns—offices can hire interns for variable lengths of time (and pay wages proportional to time worked) to assist with the production of legislation. This generates the price of a bill, p.[51]

Consider two offices A, B which employ the set of staffers $\mathcal{A}$ and $\mathcal{B}$ respectively. Let w_a denote the wage of staffer a. Then a Pareto improvement between the offices is possible if there exists $a \in \mathcal{A}$ and $b \in \mathcal{B}$ such that the following holds:

$$\left[\alpha_a^{\rho_{\text{staff}}} + \sum_{a' \in \mathcal{A}, a' \neq a} \alpha_{a'}^{\rho_{\text{staff}}}]^{\rho_{\text{rep}}/\rho_{\text{staff}}} + \gamma_A^{\rho_{\text{rep}}}\right]^{1/\rho_{\text{rep}}} - \frac{w_a}{p} \leq [[\alpha_b^{\rho_{\text{staff}}} + \sum_{a' \in \mathcal{A}, a' \neq a} \alpha_{a'}^{\rho_{\text{staff}}}]^{\rho_{\text{rep}}/\rho_{\text{staff}}} + \gamma_A^{\rho_{\text{rep}}}]^{1/\rho_{\text{rep}}} - \frac{w_b}{p}$$

$$\left[\alpha_b^{\rho_{\text{staff}}} + \sum_{b' \in \mathcal{B}, b' \neq b} \alpha_{b'}^{\rho_{\text{staff}}}]^{\rho_{\text{rep}}/\rho_{\text{staff}}} + \gamma_B^{\rho_{\text{rep}}}\right]^{1/\rho_{\text{rep}}} - \frac{w_b}{p} \leq [[\alpha_a^{\rho_{\text{staff}}} + \sum_{b' \in \mathcal{B}, b' \neq b} \alpha_{b'}^{\rho_{\text{staff}}}]^{\rho_{\text{rep}}/\rho_{\text{staff}}} + \gamma_B^{\rho_{\text{rep}}}]^{1/\rho_{\text{rep}}} - \frac{w_a}{p}$$

with at least one inequality strict. We test for all possible Pareto improving trades under the assumption that $p = 50,000$ ($50,000 of intern time produces an additional bill), restricting attention to only trades of legislative staff between offices.

Out of roughly 10 million potential trades of staffers between offices, we find that there are only 704 trades that lead to Pareto improvements. This is quite a small number of potential improvements, suggesting that offices hire (legislative) staffers with some degree of sophistication. Both the team's productivity and the Representative's productivity must be accurately incorporated into a hiring decision for this result to hold. The relatively small number of Pareto improvements also suggests that the outcome we examine, the bills produced, is not an unreasonable proxy for a Representative's (legislative) objective function.

6.5 Ideology

We have demonstrated that staffers are important for the production of legislation, but one may naturally wonder whether they matter for the ideological content of bills. We conjecture that the ideological content of the bills produced by a Congressional office is the weighted average of all individual's ideological ideal points within the office:

$$\iota_{jt} = \frac{\sum_i \omega_i \iota_i m_{i,j,t} + \omega_j \iota_j}{\sum_i \omega_i m_{i,j,t} + \omega_j} + \varepsilon_{jt} \tag{10}$$

where ω is a weight and ι is the staffer or Representative's ideal point. To estimate the equation above, we compute the average DW-Nominate score of bills produced by an

[51] Unlike staffers, intern skill is typically ex-ante unobserved. Assuming that intern skill is drawn from a common distribution across offices thus supports a common price for bills p, unlike staffer wages which may individually vary.

office in each quarter (normalized within the party), and use this as the outcome ι_{jt}. We take our estimates of the staffer and Representative effects under the baseline Bayesian model as weights, so that $\omega_i = \alpha_i, \omega_j = \gamma_j \forall i, j$—the assumption being that the producer of a bill can control the ideological content of the bill that they produce.

We present estimates of the Representative share of differences along the first and second dimensions of DW-Nominate in Figure 9. The first dimension of DW-Nominate corresponds to the typical liberal/conservative or left/right ideological divide in the US, while the second dimension has proven more difficult to characterize—scholars have hypothesized that it captures a blend of policy issues that may not fall neatly along traditional ideological lines. We find that Representatives are responsible for approximately all of the (within-party) differences in ideology along the first DW-Nominate dimension, leaving little room for staffers to influence ideology along typical partisan lines. If anything, the posterior mean for the Representative share being 122% suggests that staffers play a (weak) moderating influence on their Representatives; in their absence, the bills produced by Congress would be 22% more extreme. However, staffers account for the lion's share of differences in ideology along the second dimension. The posterior mean for the Representative share along the second dimension is a mere 16%, implying a staffer share over 80%. Staffers exert much more influence when focusing on these alternate, less partisan issues. To verify the results of our estimation process, we regress the posterior means of a Representative's estimated ideology when writing bills against their DW-Nominate score from roll call votes. Table A.6 shows that these measures tend to be significantly correlated. Taken together, these results suggest that staffers are most impactful when they pursue policy goals outside of the typical partisan debate, and that they may act as a moderating force in partisan politics.

7 Conclusion

In this paper, we have laid out a new framework for estimating the impact of individuals working within teams when individual contributions are unobserved. Through linear decompositions, dynamic estimation, and Bayesian methods, we have quantified the crucial role that Congressional staffers play in the legislative process. Our findings indicate that staffers are not merely extensions of their members of Congress but are influential contributors to legislative productivity, accounting for 40% or more of the variation in productivity between different Congressional offices. In extensions to our baseline model, we further characterize the importance of staffers, identifying the legislative positions in which they have the greatest impact, the complementarity between staffer and Represen-

tative productivity, and the ways in which they can moderate ideological extremism and contribute to non-partisan legislation.

Looking forward, there are many interesting questions to explore. First, the influence of staffers likely reaches beyond legislative productivity to other areas such as constituent services, media communication, and government oversight, as hinted in decompositions of Congressional office spending (see Figure A.1). What is the impact of Congressional staffers in these domains? Future research may investigate this question. Second, although this paper has focused on the impact of personal staff, committee staff also play a large role in the production of legislation. It would be interesting to understand the role that other staffers, and other non-elected members of government more broadly, play in the legislative process. Third, this paper raises the question of whether staffers are able to reduce polarization and make progress on non-partisan issues precisely because they are non-elected. The role of electoral accountability on political polarization more broadly speaking could be an important avenue for further research.

To conclude, we present one final speculative regression on the role that Congressional staffers play in our political system. Representatives in the House were first allowed to hire a single staffer in 1893. Over the course of the next century, Congress expanded the maximum number of staffers that Representatives could hire fifteen times. In Figure 10, we present coefficients from a two-way fixed effect event study, regressing the maximum number of staffers that Representatives can hire on the probability that the incumbent wins an election. We find that a doubling in the number of staffers corresponds to a roughly 20% increase in the probability that an incumbent wins. It appears that staffers even influence the thing that politicians may care about most: re-election.

References

Aberbach, Joel D, Robert D Putnam, and Bert A Rockman, *Bureaucrats and politicians in western democracies*, Harvard university press, 1981.

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

_ , _ , and Simon Woodcock, "Econometric analyses of linked employer–employee data," in "The econometrics of panel data: Fundamentals and recent developments in theory and practice," Springer, 2008, pp. 727–760.

_ , Kevin L McKinney, and Ian M Schmutte, "Modeling endogenous mobility in wage determiniation," *US Census Bureau Center for Economic Studies Paper No. CES-WP-15-18*, 2015.

Ahmadpoor, Mohammad and Benjamin F Jones, "Decoding team and individual impact in science and invention," *Proceedings of the National Academy of Sciences*, 2019, *116* (28), 13885–13890.

Andrews, Martyn J, Len Gill, Thorsten Schank, and Richard Upward, "High wage workers and low wage firms: negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society Series A: Statistics in Society*, 2008, *171* (3), 673–697.

Aneja, Abhay and Guo Xu, "Strengthening State Capacity: Civil Service Reform and Public Sector Performance during the Gilded Age," *The American Economic Review*, Forthcoming.

Angrist, Joshua D, Peter D Hull, Parag A Pathak, and Christopher R Walters, "Leveraging lotteries for school value-added: Testing and estimation," *The Quarterly Journal of Economics*, 2017, *132* (2), 871–919.

Arora, Abhishek and Melissa Dell, "LinkTransformer: A Unified Package for Record Linkage with Transformer Language Models," *arXiv preprint arXiv:2309.00789*, 2023.

Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter, "Management practices, workforce selection, and productivity," *Journal of Labor Economics*, 2018, *36* (S1), S371–S409.

Berger, James O, *Statistical decision theory and Bayesian analysis*, Springer Science & Business Media, 2013.

Bergeron, Augustin, Pedro Bessone, John Kabeya Kabeya, Gabriel Z Tourek, and Jonathan L Weigel, "Optimal assignment of bureaucrats: Evidence from randomly assigned tax collectors in the DRC," Technical Report, National Bureau of Economic Research 2022.

Bertrand, Marianne, Matilde Bombardini, and Francesco Trebbi, "Is It Whom You Know or What You Know? An Empirical Assessment of the Lobbying Process," *American Economic Review*, December 2014, *104* (12), 3885–3920.

Besley, Timothy and Torsten Persson, "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, August 2009, *99* (4), 1218–1244.

Boche, Adam, Jeffrey B Lewis, Aaron Rudkin, and Luke Sonnet, "The new Voteview. com: preserving and continuing Keith Poole's infrastructure for scholars, students and observers of Congress," *Public Choice*, 2018, *176*, 17–32.

Bonhomme, Stéphane, "Teams: Heterogeneity, sorting, and complementarity," *Proceedings of the 15th World Congress of the Econometric Society*, 2021.

_ , Thibaut Lamadon, and Elena Manresa, "A distributional framework for matched employer employee data," *Econometrica*, 2019, *87* (3), 699–739.

Cantoni, Enrico and Vincent Pons, "Does context outweigh individual characteristics in driving voting behavior? Evidence from relocations within the United States," *American Economic Review*, 2022, *112* (4), 1226–1272.

Card, David, Jörg Heining, and Patrick Kline, "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly journal of economics*, 2013, *128* (3), 967–1015.

Chan, David C, "Influence and information in team decisions: evidence from medical residency," *American Economic Journal: Economic Policy*, 2021, *13* (1), 106–137.

Chen, Yiqun, "Team-specific human capital and team performance: evidence from doctors," *American economic review*, 2021, *111* (12), 3923–3962.

Chetty, Raj and Nathaniel Hendren, "The impacts of neighborhoods on intergenerational mobility II: County-level estimates," *The Quarterly Journal of Economics*, 2018, *133* (3), 1163–1228.

_ , John N Friedman, and Jonah E Rockoff, "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates," *American economic review*, 2014, *104* (9), 2593–2632.

Childers, David, Jesús Fernández-Villaverde, Jesse Perla, Christopher Rackauckas, and Peifan Wu, "Differentiable state-space models and hamiltonian monte carlo estimation," Technical Report, National Bureau of Economic Research 2022.

Constantine, Noah and Sergio Correia, "Putting the "i" in Fixed Effects: Individual Fixed Effects with Group Level Outcomes," Technical Report, Working paper 2021.

Cross, Jesse M and Abbe R Gluck, "The Congressional Bureaucracy," *University of Pennsylvania Law Review*, 2020, *168* (6), 1541–1683.

Crosson, Jesse M, Alexander C Furnas, Timothy Lapira, and Casey Burgat, "Partisan competition and the decline in legislative capacity among congressional offices," *Legislative Studies Quarterly*, 2021, *46* (3), 745–789.

_ , Geoffrey M Lorenz, Craig Volden, and Alan E Wiseman, "How experienced legislative staff contribute to effective lawmaking," in "Congress overwhelmed: The decline in congressional capacity and prospects for reform" University of Chicago Press Chicago, IL 2020, pp. 209–24.

Dahlström, Carl and Victor Lapuente, "Comparative bureaucratic politics," *Annual Review of Political Science*, 2022, *25*, 43–63.

DeGregorio, Christine, "Professionals in the US Congress: an analysis of working styles," *Legislative Studies Quarterly*, 1988, pp. 459–476.

Dikkala, Nishanth, Greg Lewis, Lester Mackey, and Vasilis Syrgkanis, "Minimax estimation of conditional moment models," *Advances in Neural Information Processing Systems*, 2020, *33*, 12248–12262.

Ferguson, Thomas S, *Mathematical statistics: A decision theoretic approach*, Vol. 1, Academic press, 1967.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams, "Place-based drivers of mortality: Evidence from migration," *American Economic Review*, 2021, *111* (8), 2697–2735.

Fox, Harrison W and Susan Webb Hammond, *Congressional staffs: The invisible force in American lawmaking*, New York: The Free Press, 1977.

Freund, Lukas, "Superstar Teams," *Available at SSRN 4312245*, 2024.

Gailmard, Sean and John W Patty, "Formal models of bureaucracy," *Annual Review of Political Science*, 2012, *15*, 353–377.

Gelman, Andrew, John B Carlin, Hal S Stern, and Donald B Rubin, *Bayesian data analysis*, Chapman and Hall/CRC, 1995.

Hagedorn, Sara L, "Taking the lead congressional staffers and their role in the policy process." PhD dissertation, University of Colorado at Boulder 2015.

Hall, Richard L and Alan V Deardorff, "Lobbying as legislative subsidy," *American Political Science Review*, 2006, *100* (1), 69–84.

Hoffman, Mitchell and Christopher T Stanton, "People, Practices, and Productivity: A Review of New Advances in Personnel Economics," Working Paper 32849, National Bureau of Economic Research August 2024.

i Vidal, Jordi Blanes, Mirko Draca, and Christian Fons-Rosen, "Revolving Door Lobbyists," *American Economic Review*, December 2012, *102* (7), 3731–48.

Iranzo, Susana, Fabiano Schivardi, and Elisa Tosetti, "Skill dispersion and firm productivity: An analysis with employer-employee matched data," *Journal of Labor Economics*, 2008, *26* (2), 247–285.

Kingdon, John W, *Congressmen's voting decisions*, University of Michigan Press, 1989.

Lazarus, Jeffrey, Joshua McCrain, and Amy McKay, "Lobbying access, congressional employment, and the determinants of revolving door lobbying," 2019.

Leal, David L and Frederick M Hess, "Who chooses experience? Examining the use of veteran staff by house freshmen," *Polity*, 2004, *36* (4), 651–664.

Lehmann, Erich Leo, Joseph P Romano, and George Casella, *Testing statistical hypotheses*, Vol. 3, Springer, 1986.

Montgomery, Jacob M and Brendan Nyhan, "The effects of congressional staff networks in the US House of Representatives," *The Journal of Politics*, 2017, *79* (3), 745–761.

Niskanen, William A, "The peculiar economics of bureaucracy," *The American Economic Review*, 1968, *58* (2), 293–305.

Ommundsen, Emily Cottle, "The Institution's Knowledge: Congressional Staff Experience and Committee Productivity," *Legislative Studies Quarterly*, 2023, *48* (2), 273–303.

Poole, Keith T and Howard Rosenthal, "A spatial model for legislative roll call analysis," *American journal of political science*, 1985, pp. 357–384.

Romzek, Barbara S. and Jennifer A. Utter, "Congressional Legislative Staff: Political Professionals or Clerks?," *American Journal of Political Science*, 1997, *41* (4), 1251–1279.

Salisbury, Robert H. and Kenneth A. Shepsle, "Congressional Staff Turnover and the Ties-That-Bind," *American Political Science Review*, 1981, *75* (2), 381–396.

Shepherd, Michael E and Hye Young You, "Exit strategy: career concerns and revolving doors in Congress," *American Political Science Review*, 2020, *114* (1), 270–284.

Silver, David, "Haste or waste? Peer pressure and productivity in the emergency department," *The Review of Economic Studies*, 2021, *88* (3), 1385–1417.

Smith, Steven S, Jason M Roberts, and Ryan J Vander Wielen, *The American Congress*, Cambridge University Press, 2013.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter, "Firming up inequality," *The Quarterly journal of economics*, 2019, *134* (1), 1–50.

Spenkuch, Jörg, Edoardo Teso, and Guo Xu, "Ideology and Performance in Public Organizations," *Econometrica*, Forthcoming.

Vogler, Jan P, "Bureaucracies in historical political economy," in "The Oxford Handbook of Historical Political Economy," Oxford University Press, 2022, pp. C19–S1.

Volden, Craig and Alan E Wiseman, "Legislative effectiveness in Congress," *Manuscript, The Ohio State University*, 2009.

Weber, Max, *Economy and society*, New York: Bedminster Press, 1968.

Weidmann, Ben, Joseph Vecci, Farah Said, David J Deming, and Sonia R Bhalotra, "How Do You Find a Good Manager?," Technical Report, National Bureau of Economic Research 2024.

Wilkerson, John, David Smith, and Nicholas Stramp, "Tracing the flow of policy ideas in legislatures: A text reuse approach," *American Journal of Political Science*, 2015, *59* (4), 943–956.

Tables and Figures

Table 1: Summary statistics

	Overall			Movers	Non-movers	t-test
	Mean	Min	Max	Mean	Mean	p: 4 = 5
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Staffer level						
Career length (quarters)	9.30	0.00	468.57	15.50	8.28	0.000
	(13.08)			(11.76)	(13.00)	
Wage	11179.44	-8958.32	52911.11	11813.41	11075.48	0.000
	(6440.21)			(5402.13)	(6589.18)	
Share career as legislative staff	0.23	0.00	1.00	0.33	0.21	0.000
	(0.39)			(0.41)	(0.39)	
Share career as managment	0.19	0.00	1.00	0.22	0.19	0.000
	(0.35)			(0.34)	(0.35)	
Share career as legislative director	0.03	0.00	1.00	0.07	0.03	0.000
	(0.15)			(0.18)	(0.15)	
Share career as chief of staff	0.04	0.00	1.00	0.05	0.04	0.005
	(0.19)			(0.20)	(0.19)	
Panel B: Staffer level (at least one year)						
Job length (quarters)	5.14	0.11	41.43	4.28	5.47	0.000
	(5.25)			(4.19)	(5.57)	
Office length (quarters)	6.89	0.11	41.43	5.63	7.40	0.000
	(6.68)			(5.35)	(7.09)	
# jobs (per office)	1.34	1.00	7.00	1.32	1.35	0.000
	(0.68)			(0.65)	(0.69)	
Ever switch job category (in office)	0.88	0.00	1.00	0.88	0.88	0.844
	(0.33)			(0.33)	(0.33)	
# offices over career	1.36	1.00	132.00	2.74	1.13	0.000
	(2.67)			(4.76)	(2.04)	

Notes: This table presents summary statistics for personal staffers in the House of represents. In Panel A, we present summary statistics for all staffers. Wages are in US dollars per quarter. In Panel B, we present summary statistics for all staffers who work in the House for at least one year. Job length is the amount of time (in quarters) that a staffer holds the same job title. Office length is the amount of time (in quarters) that a staffer works for the same Representative. Standard deviations presented below means in parenthesis.

Table 2: Correlating staffer FEs with wages

FE source:	Bills cospons.	Laws passed	Exit comm.	Pass House	Important bills	Index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: staffer level						
abs(Staffer FE)	0.080***	0.056**	0.114***	0.119***	0.042**	0.084***
	(0.023)	(0.022)	(0.025)	(0.025)	(0.018)	(0.018)
Panel B: staffer level, only those with unique work histories						
abs(Staffer FE)	0.087***	0.065***	0.124***	0.127***	0.042**	0.091***
	(0.024)	(0.022)	(0.026)	(0.026)	(0.018)	(0.018)
Panel C: history level						
abs(Staffer FE)	0.084***	0.062***	0.121***	0.125***	0.042**	0.089***
	(0.023)	(0.022)	(0.025)	(0.025)	(0.018)	(0.018)
Panel D: history level, staffer weighted						
abs(Staffer FE)	0.079***	0.055**	0.114***	0.118***	0.041**	0.083***
	(0.024)	(0.022)	(0.025)	(0.025)	(0.018)	(0.018)
Panel E: history level, managers only						
abs(Staffer FE)	0.124**	0.203***	0.327***	0.318***	0.097*	0.221***
	(0.056)	(0.065)	(0.055)	(0.056)	(0.054)	(0.047)
Panel F: history level, managers and legislative staff only						
abs(Staffer FE)	0.099***	0.101***	0.169***	0.170***	0.048	0.123***
	(0.033)	(0.033)	(0.036)	(0.036)	(0.031)	(0.027)

Notes: Regressions are at the staffer level in Panels A-B, and at the work history level in Panels C-F (independent and dependent variables for all staffers that are always in the same office at the same time are averaged to a single observation). In Panel B, we restrict the sample to staffers who have a unique work history. In Panel D, observations are weighted by the number of staffers who share that work history. In Panel E, we restrict the sample to staffers who work in a management position. In Panel F, we restrict the sample to staffers who work in a management or legislative position. Independent and dependent variables of interest have been standardized to mean 0, and standard deviation 1. Staffer fixed effects α_i are estimated from the linear team model in Equation 1. Each column presents an estimate using a different outcome variable, and column 6 is an index consisting of the sum of the prior five columns, standardized. Wages are a staffer's average wage over their career. Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

Table 3: Bayesian model

Param	Mean	2.5%	97.5%	$\tilde{n}$	$\hat{R}$
(1)	(2)	(3)	(4)	(5)	(6)
Panel A: baseline model					
μ_α	-5.38	-7.20	-4.03	680	1.01
μ_γ	2.52	2.43	2.60	647	1.00
σ_α	3.38	2.77	4.18	545	1.01
σ_γ	0.52	0.48	0.57	405	1.00
σ_τ	12.80	9.91	16.42	186	1.03
σ	10.22	10.08	10.38	2361	1.00
Rep. share	0.63	0.59	0.66	1071	1.00
Panel B: CES model					
ρ_{staff}	0.79	0.62	0.89	4014	1.00
ρ_{rep}	0.13	0.10	0.17	169	1.03
μ_α	-39.96	-51.91	-29.22	1817	1.01
μ_γ	2.36	2.24	2.47	450	1.02
σ_α	9.98	7.50	12.91	1680	1.00
σ_γ	0.54	0.50	0.60	459	1.01
σ_τ	12.65	9.60	16.75	138	1.04
σ	10.05	9.91	10.19	5348	1.00
Rep. share	0.57	0.53	0.63	1133	1.00

Notes: This table presents parameter estimates from the baseline model (Panel A) and CES model (Panel B). Columns 2 through 4 display posterior means, 2.5th percentiles, and 97.5th percentiles respectively. Column 5 displays the effective sample size and column 6 displays the $\hat{R}$, a measure of MCMC convergence. $\mu_\alpha, \sigma_\alpha$ are the staffer mean and standard deviation, $\mu_\gamma, \sigma_\gamma$ are the Representative mean and standard deviation. Both parameterize log-normal distributions. σ_τ is the time standard deviation, and σ is the error standard deviation. Rep. share is the share of differences between above and below median bill output offices that is attributable to Representatives. $\rho_{\text{staff}}, \rho_{\text{rep}}$ are elasticities of substitution within a team of staffers, and between the staffers and Representative, respectively.

Table 4: Variance decomposition

	Share	95% CI
	(1)	(2)
Staffer share	37.04%	[33.11%, 40.87%]
Office share	85.86%	[82.77%, 88.90%]

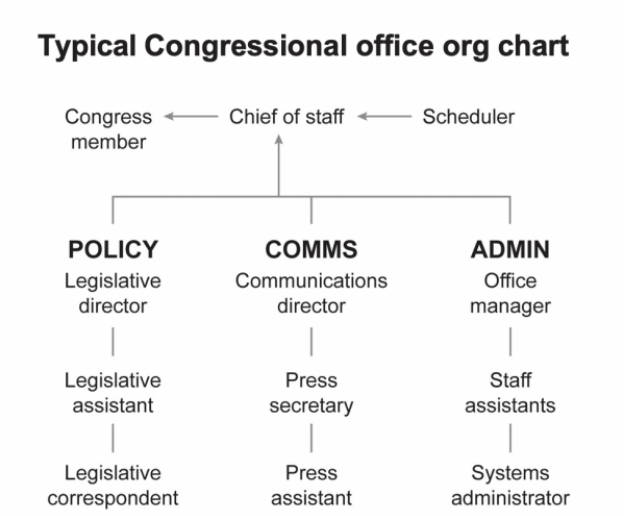
Notes: Variance decomposition from Bayesian estimates of Equation 1. Column (1) presents staffer and office shares of variance at the office level, and column (2) presents 95% credible intervals.

Table 5: Effect of diversity on bills passed

	Bills passed			
	(1)	(2)	(3)	(4)
Race HHI	-0.681			
	(0.025)			
Gender HHI		-0.975		
		(0.045)		
Education HHI			-0.591	
			(0.023)	
Work experience HHI				-1.346
				(0.109)
95% CI	[-0.729, -0.633]	[-1.065, -0.891]	[-0.636, -0.547]	[-1.558, -1.133]

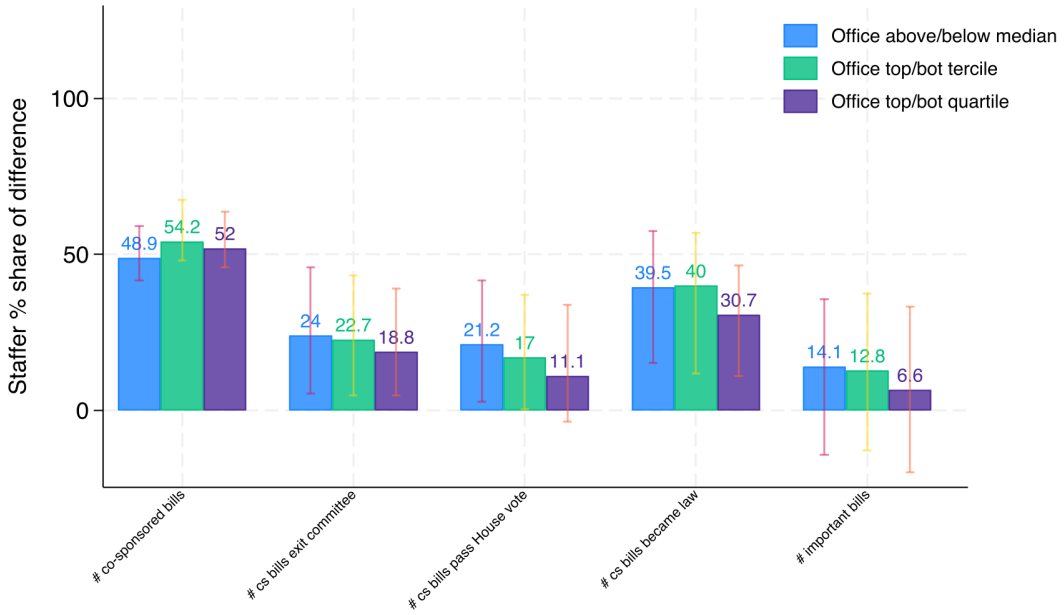
Notes: This table presents coefficients and standard errors from a Bayesian estimate of Equation 1 that additionally includes the indicated covariate. HHI is the sum of squared shares within the office for a given quarter. Column (1) presents results for different ethnicities, column (2) for gender, column (3) for the highest educational degree attained, and column (4) for types of work experience prior to Congress (private sector, policy-facing, state government, and federal government). 95% credible intervals are presented in the final row.

Figure 1: Staffer org chart



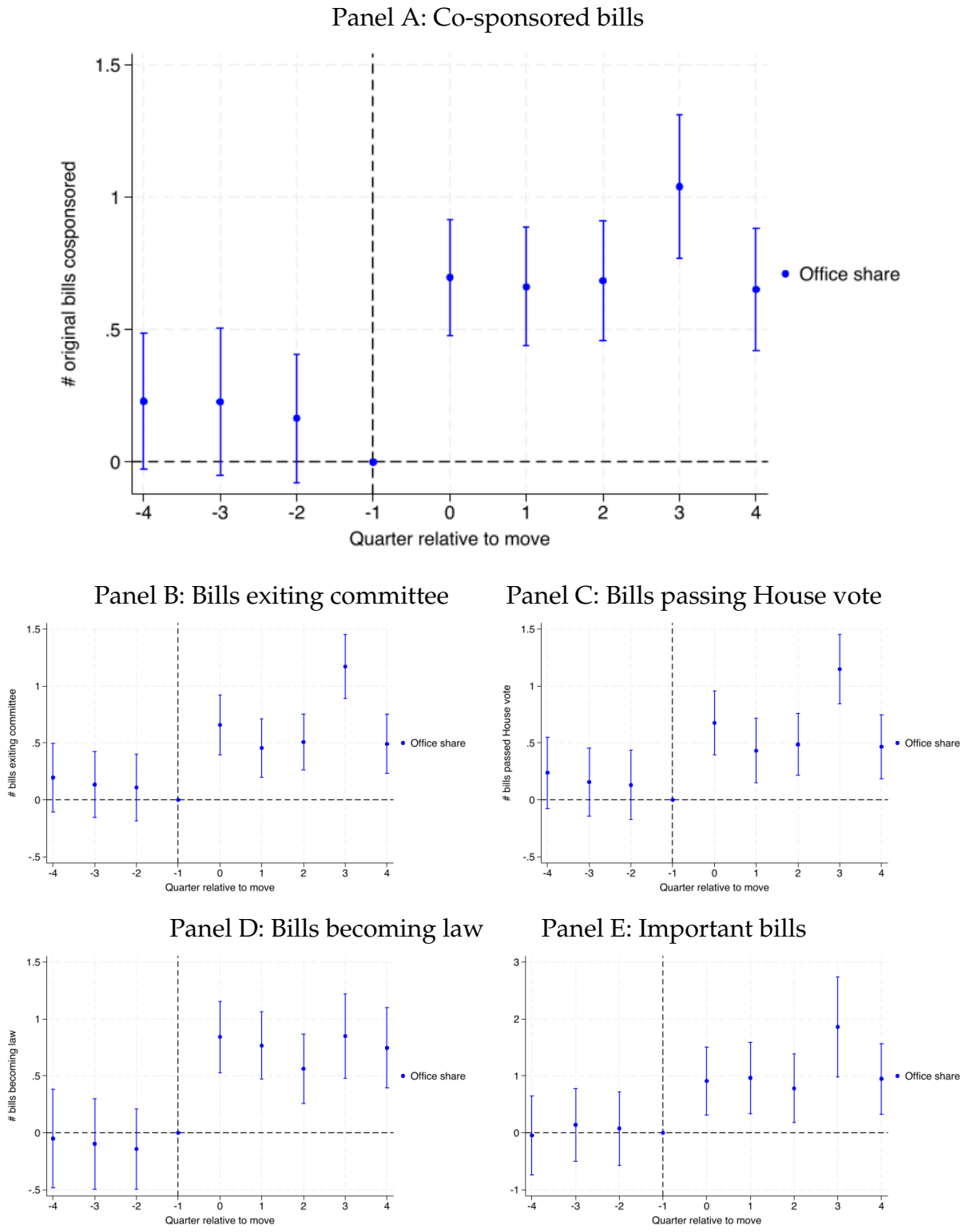
Note: Source `https://www.politicopro.com/blog/congressional-office-org-chart/`

Figure 2: Staffer share of differences in legislative productivity, linear decomposition



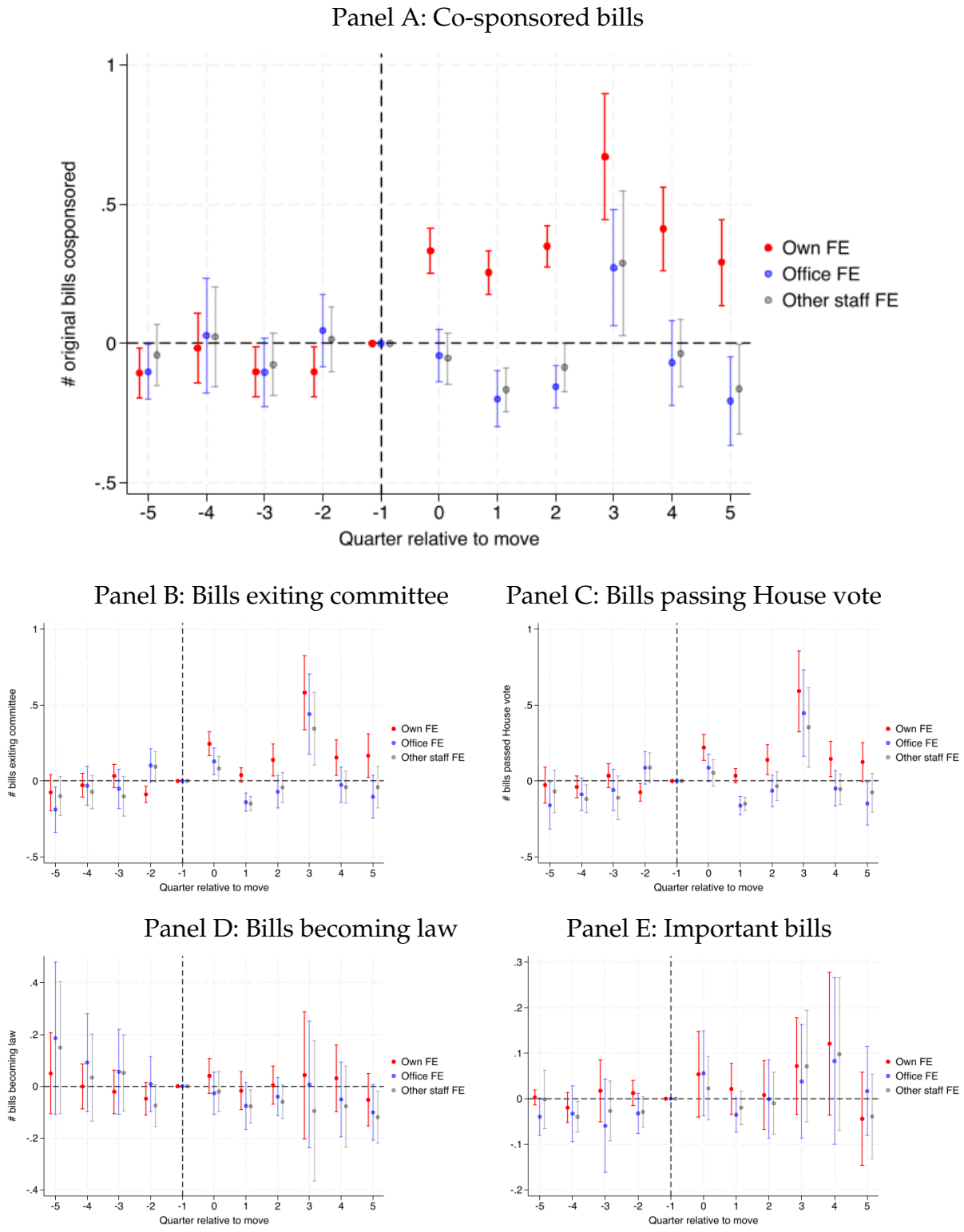
Note: This figure plots the staffer share of differences between offices that are above and below median in legislative output in blue, as well as 95% confidence intervals. The staffer share is computed from fixed effects estimated in a team-based mover design framework (Equation 1). Staffer shares for the difference in top and bottom tercile (quartile) offices are plotted in green (purple). Bars are arranged from left to right in order of closeness to passage into law, except for 'important bills', which refers to the number of bills referenced in the Congressional Quarterly Almanac year-end write-up. 'cs bills' refers to the number of bills co-sponsored by the office.

Figure 3: Dynamic estimation of Representative share of differences in legislative productivity



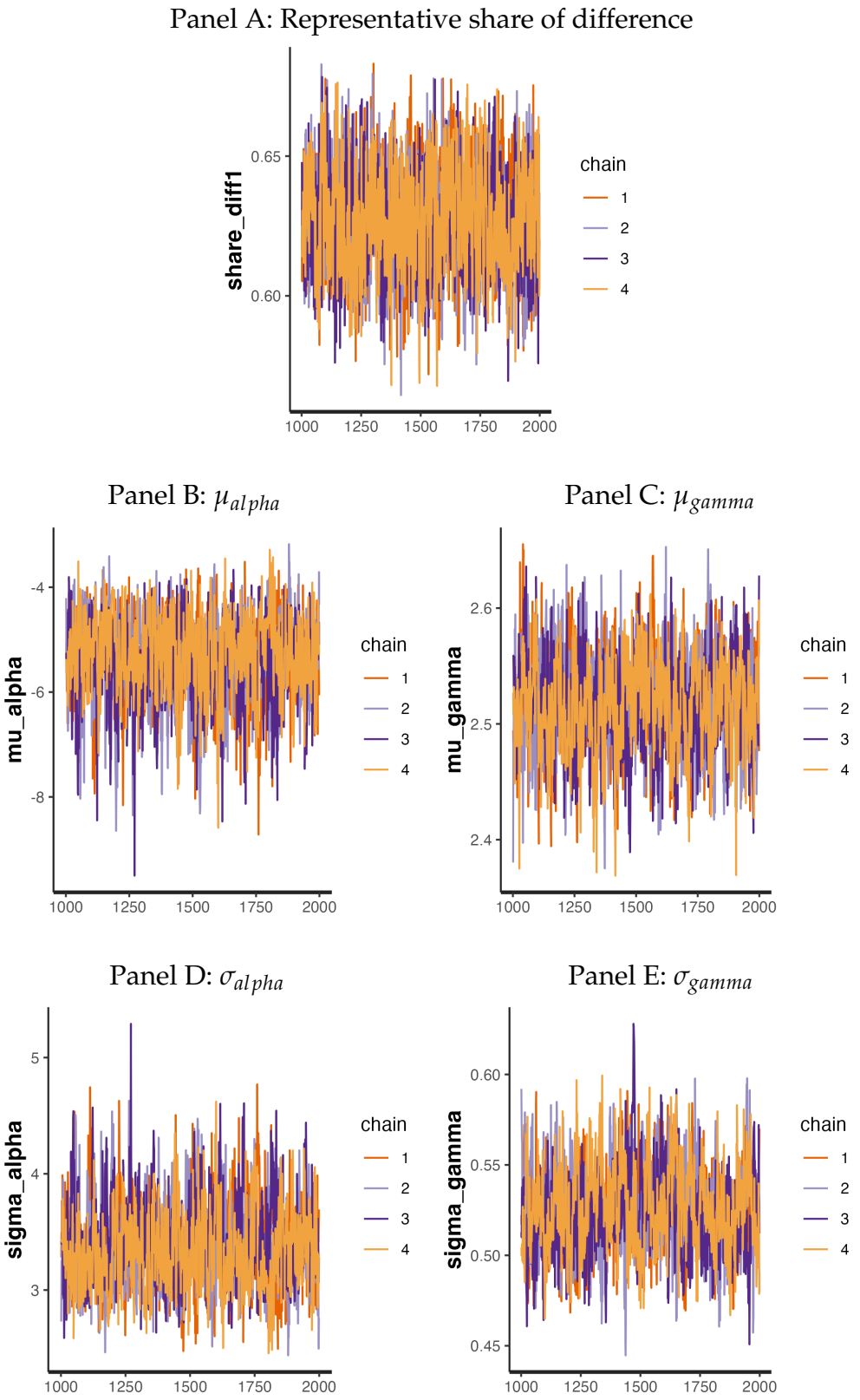
Note: This figure plots coefficients and 95% confidence intervals for an event study specification (Equation 5) at the staffer-quarter level (for even quarters only). The x-axis shows quarters to a staffer's first move. Displayed coefficients are the Representative share of differences between the destination and origin office's legislative output. The specific type of legislative output varies by panel. Fixed effects for other staffers on the team (estimated out of sample on odd quarters) are subtracted from the outcome. Standard errors are clustered at the Representative level.

Figure 4: Dynamic prediction of staffer movement using fixed effects



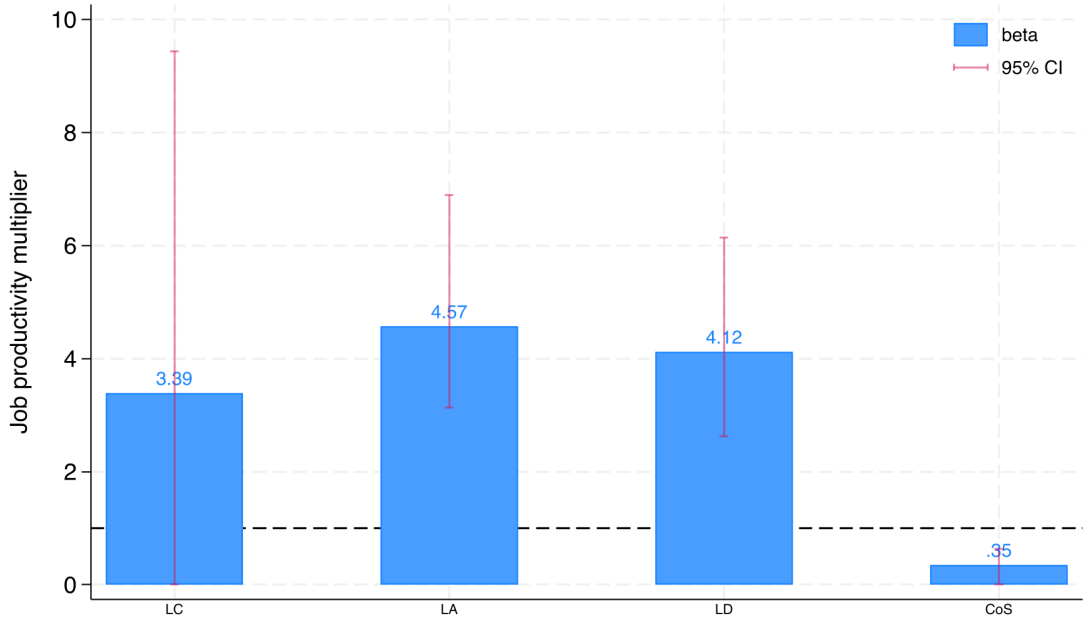
Note: This figure plots coefficients and 95% confidence intervals for an event study specification (Equation 6) at the staffer-quarter level (for even quarters only). The x-axis shows quarters to a staffer's first move. The outcome is a specific type of destination office legislative output that varies by panel. Coefficients on the moving staffer's fixed effect are plotted in red, coefficients on the destination Representative's fixed effect are plotted in blue, and coefficients on the destination team of staffers' fixed effect are plotted in gray. These fixed effects are estimated out of sample on odd quarters. Standard errors are clustered at the Representative level.

Figure 5: Trace plots for the baseline model, number of bills co-sponsored



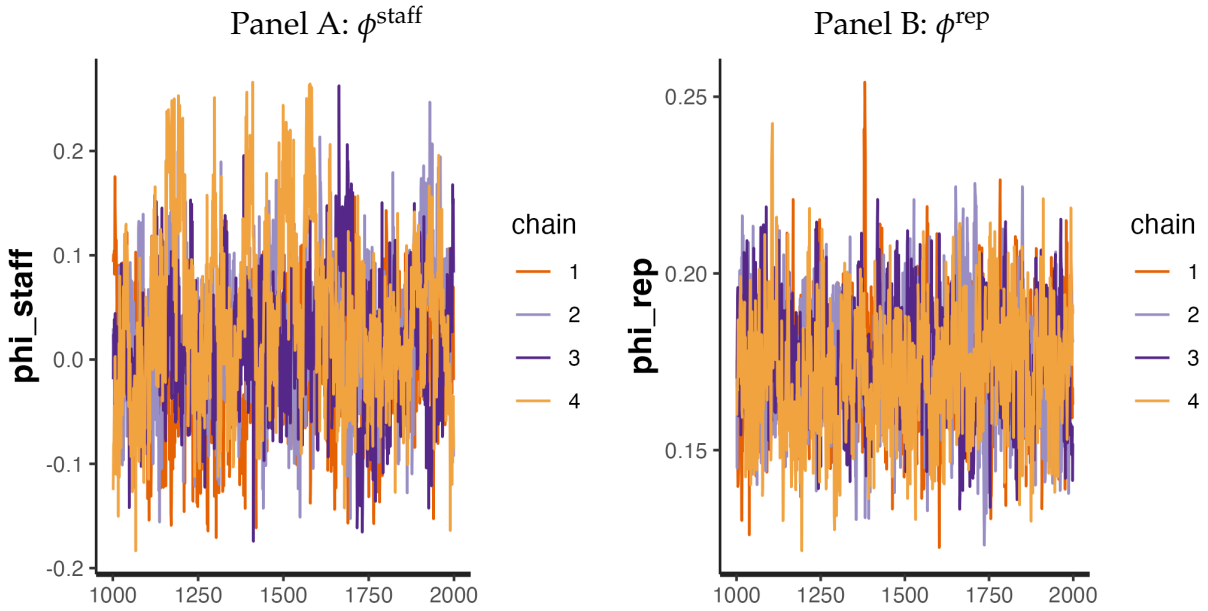
Note: This figure presents trace plots for a Bayesian estimation of the team-based mover design framework (Equation 1). The x-axis represents samples from a posterior drawn via MCMC. Panel A shows the Representative share of differences between above and below-median offices in bills produced. Panels B and D (C and E) display the staffer (Representative) log-normal distribution's mean and standard deviation. Values for 4 different chains are displayed.

Figure 6: Job β productivities



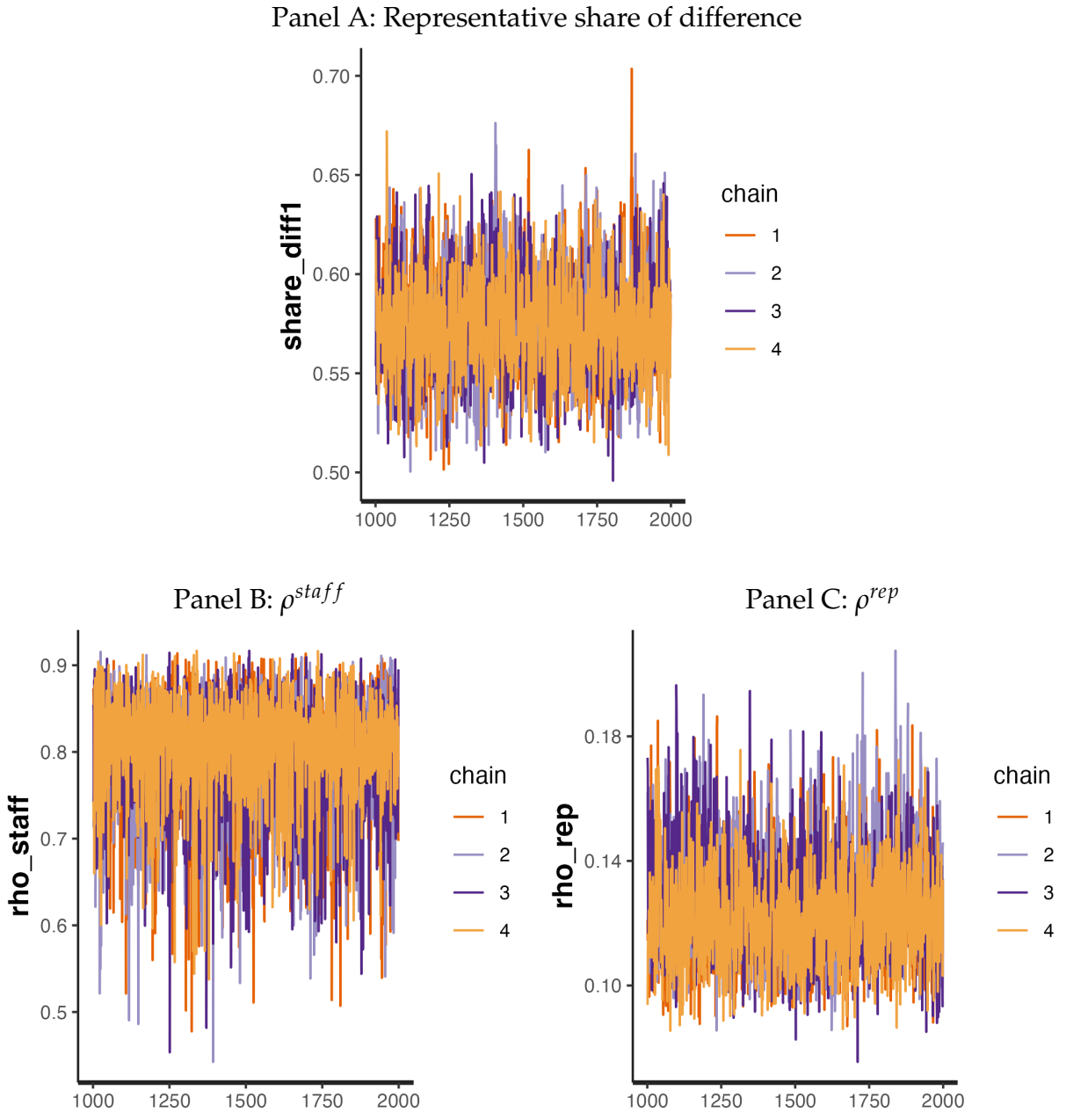
Note: This figure presents posterior means and 95% confidence intervals for job-specific slopes on staffer fixed effects (see Equation 7). LC: legislative correspondent, LA: legislative assistant/aide, LD: legislative director, CoS: chief of staff. All other jobs are normalized to have slope 1.

Figure 7: Trace plots for experience model



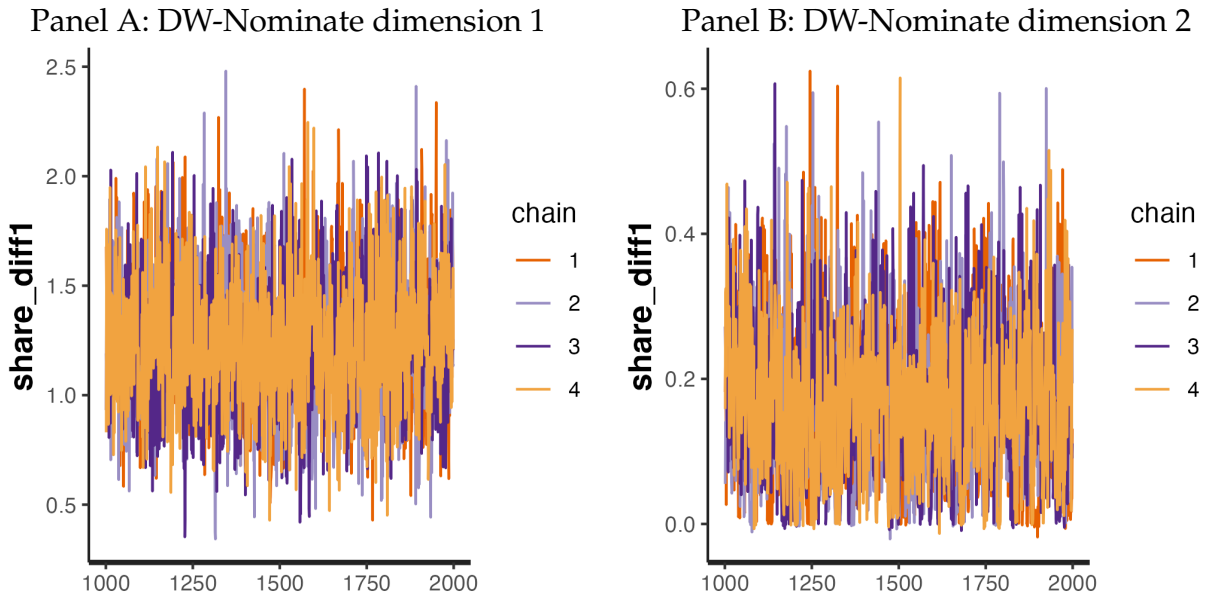
Note: This figure presents trace plots for a Bayesian estimation of the team-based mover design framework featuring returns to time spent in Congress (Equation 8). The x-axis represents samples from a posterior drawn via MCMC. Panel A shows the return for staffers and Panel B shows the return for Representatives. Values above 0 indicate positive returns to experience. Values for 4 different chains are displayed.

Figure 8: Trace plots for CES model



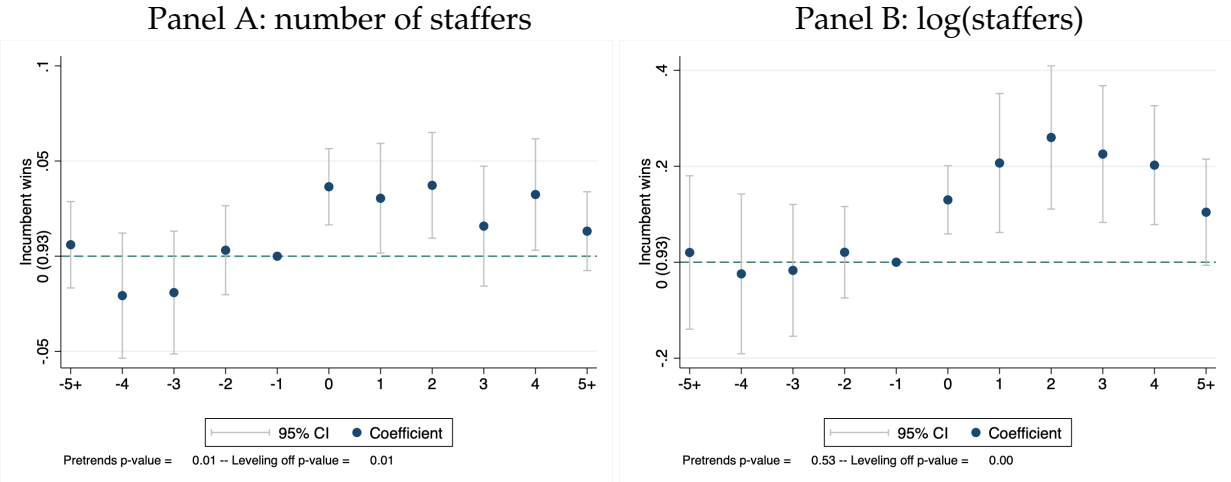
Note: This figure presents trace plots for a Bayesian estimation of the team-based mover design framework featuring a nested CES production function (Equation 9). The x-axis represents samples from a posterior drawn via MCMC. Panel A shows the Representative share of differences between above and below median offices in bills produced. Panel B presents the elasticity of substitution between staffers working on the same team, while Panel C presents the elasticity of substitution between the team of staffers and the Representative. Values for 4 different chains are displayed.

Figure 9: Trace plots for ideology model



Note: This figure presents trace plots for a Bayesian estimation of the ideology of bills produced by an office (Equation 10), which is a weighted average of Representative and staffer ideal points. Weights are estimated from the team-based mover design (Equation 1). The x-axis represents samples from a posterior drawn via MCMC. Both panels show the Representative share of differences between above and below-median offices in ideology. Panel A estimates ideology for the first DW-Nominate dimension, the conventional left-right partisan ideology. Panel B estimates ideology for the second DW-Nominate dimension, concentrated on other policy domains. Values for 4 different chains are displayed.

Figure 10: The effect of additional staffers on the incumbency advantage



Note: This figure presents coefficients and 95% confidence intervals from a two-way fixed effect event study design. Each event is an expansion in the number of staffers that Representatives in the House were legally allowed to hire (of which there were 15 total). The regression is at the Congressional district election year level, with Senators included for the estimation of time-fixed effects. The outcome is an indicator of the incumbent winning an election. Panel A plots the coefficient for the maximum number of staffers allowed to be hired, while Panel B plots the coefficient for the log of the maximum number of staffers allowed to be hired. Standard errors are clustered at the Congressional district level.

Online Appendix for: Puppetmasters or Pawns? The Power of Congressional Staffers

This appendix contains additional figures and tables for the article "Puppetmasters or Pawns? The Power of Congressional Staffers."

Appendix A Data

Constructing a panel of staffers We construct panel data for staffers, associating each staffer with a single office in each quarter, using the following procedure:

1. We drop all staffers that never move.
2. We drop all staffers who are interns or part-time workers.
3. We keep only staffers that stay in Congress for at least one year (4 quarters).
4. We drop any staffers that work for more than 5 offices in a single quarter. These tend to be staffers in generalist positions (such as technology consultants) who are outside the main focus of our study.
5. For each staffer that works for multiple offices in a quarter, we assign an office to a staffer based on the longest duration worked. We break ties randomly.
6. If a staffer has spent less than 40% of a quarter working for the office, we also drop them.
7. We identify a job as a continuous period of time spent working for the same office, with up to 1 quarter of interruption in the middle.

Important bills and the LES We quote the Center for Effective Lawmaking on their procedure for constructing important bills (which they term "substantive and significant"):

> A bill is deemed substantive and significant if it was introduced during the 93rd – 113th Congresses, and had been the subject of an end-of-the-year write-up in the Congressional Quarterly Almanac. For bills that were introduced into the 114th to 116th Congresses, a bill is deemed substantive and significant if it was mentioned on two or more occasions in the stories published in Congressional Quarterly Weekly/CQ Magazine during that Congress. For bills that were introduced into the 117th Congress and subsequent congresses, a bill is deemed substantive and significant if it was mentioned on two or more occasions in the stories published in Congressional Quarterly Weekly/CQ Magazine during that Congress, identified as the subject of a key vote by Project Vote Smart, and/or discussed in the Fact Sheets in the CQ House section of CQ.com.

Many important bills have large numbers of co-sponsors (often in the dozens). Given that we have no effective way to determine which offices were primarily responsible for crafting the original text of these bills, we drop all important bills that received five or more original co-sponsors.

The Legislative Effectiveness Score (LES) is constructed as a weighted average of bills introduced, bills receiving committee action, bills passing the house, and bills becoming

law, weighted by their importance. The benchmark is the predicted LES from an OLS regression containing covariates Representative seniority and dummies for party majority, and committee and subcommittee chair.[1]

Measures of diversity To assess diversity at the office level, we develop a set of diversity measures across race, gender, educational background, and work experience using the Herfindahl-Hirschman Index (HHI). We begin by gathering staffer-level data from LegiStorm, a comprehensive database of congressional staffers. We use staffer names from the Congressional Statements of Disbursement to search for corresponding profiles on LegiStorm. Staffers who have profiles on LegiStorm are considered successfully matched with the names from the Congressional Statements of Disbursement. LegiStorm provides staffer information including gender, work history, and degrees attained, all of which can be systematically scraped for each individual. This allows us to construct detailed profiles for our analysis.

We construct data on race by employing an open-source tool called ethnicolr[2], which predicts race and ethnicity based on staffer names. This tool is trained on census names and is useful when race information is not available.

To quantify educational diversity, we parse the educational backgrounds listed on LegiStorm profiles. By identifying and categorizing keywords such as "BS", "BA", "Bachelor", and equivalent terms, we classify the degrees attained by each staffer. Degrees are categorized into levels such as bachelor's, master's, MBA, MD, JD, and PhD. This classification enables us to gauge the range and concentration of educational achievements within an office.

In terms of work experience, we leverage the capabilities of GPT-4 to automate and refine the categorization process. We first define categories for work experiences, including internship, federal government, state government, policy, and private sector roles. We then input the entire work history of staffers, focusing on experiences prior to their first congressional role (excluding congressional internships), and use GPT-4 to classify these experiences according to the predefined categories. GPT-4 is particularly useful with this kind of task due to its extensive training on billions of texts across diverse domains. This vast training enables GPT-4 to develop a deep understanding of language patterns, context, and domain-specific knowledge, making it highly effective at tasks like text classification. One concern with using language models like GPT-4 is the potential for "hallucination," where the model might generate or infer information that isn't present in the data. However, in this specific use case, GPT-4 is effective for the following reasons:

1. GPT-4 is provided with specific work histories, which are real, structured inputs. It is tasked with categorizing this information into predefined categories (e.g., federal government, private sector). The process doesn't involve generating new content or making speculative interpretations; instead, it relies on its ability to recognize patterns within the existing data.

2. The task is narrow and well-defined: to classify work experience into categories.

[1]The specific algorithm used is outlined in `https://thelawmakers.org/methodology`.

[2]More details of this package can be found at `https://github.com/appeler/ethnicolr`.

This specificity reduces the likelihood of hallucination, as GPT-4 isn't asked to extrapolate beyond the information provided. It's simply applying its understanding to categorize text, which is a core strength of language models when given clear parameters.

3. GPT-4 excels at understanding context, which is critical when distinguishing between different types of work experiences. For example, if the input describes a role as a "Legislative Analyst," GPT-4 uses the surrounding context within the work history to categorize this role accurately. This ability to leverage context helps ensure that the classification is based on the actual data, not on fabricated information.

4. We also implement a human-in-the-loop approach for validation. After GPT-4 categorizes the work experiences, we go through a review process that ensures the classifications are accurate. This further mitigates the risk of potential errors or hallucinations, ensuring that the final dataset is reliable.

Finally, we compute the HHI for each of the measures described above at the office-quarter level.

Appendix B Estimation

Appendix B.1 Linear decomposition

Main model Under the typical mover design model, the number of fixed effects is constant per observation (oftentimes three—one for the individual, one for the group, and one for time). Common fixed effect estimation routines are able to quickly compute the full set of fixed effects through a variety of optimization techniques. However, under the team-based mover design model, the number of fixed effects to estimate varies per observation: a team with eleven staffers has five more fixed effects than a team with six staffers. Although recent progress has been made in estimation when these fixed effects are incidental parameters (Constantine and Correia, 2021), the recovery of the actual fixed effects in a computationally efficient way is an open problem. As a result, to estimate the fixed effects in Equation 1, we simply use the full design matrix containing indicators for all staffers and Representatives when estimating this model, applying Frisch-Waugh Lowell to residualize any time fixed effects and other incidental parameters.

After applying the standard OLS matrix algebra, we compute the Representative share of differences between high and low legislative output offices and subtract it from 1 to obtain the staffer share. We use the Bayesian bootstrap to compute confidence intervals, weighting each observation by the square root of a random variable distributed $\exp(1)$. Throughout the paper, we use 100 iterations of the bootstrap.

Instrumental variables Our instrumental variable strategy uses two sets of instruments for staffer moves at time t:

1. Turnover: This is an indicator equal to 1 if the origin office lost their election at $t-1$ and the destination office is in the same state as the origin office.

2. Migration push/pull interaction: This is the fully saturated interaction between a push factor out of an origin office into a destination (the leave-out total number of staffers who ever move between origin and destination) and a pull factor into the destination (the number of staffers who leave the office at time t).

Both of these instruments are arguably uncorrelated with the moving staffer's productivity, potentially satisfying the exclusion restriction. One may be concerned that the turnover instrument may directly influence legislative productivity. We argue that this is unlikely to be substantial for multiple reasons: first, the Representative fixed effect captures any persistent differences between Representatives who win versus lose re-election. Second, the Representative that loses an election at $t-1$ is no longer in the data at time t, so a violation of the exclusion restriction would have to operate through an anticipated loss affecting legislative productivity prior to the election. However, the incumbent advantage in the US Congress is large, mitigating this potential channel.

We begin by running a standard (linear) first stage, at the staffer-destination office-quarter level. Results are displayed in Table A.3. We note that the most predictive covariate is, unsurprisingly, the destination and origin office being the same. The instruments are all positive and statistically significant, working in the direction that we would expect. Results are robust to the inclusion of various fixed effects, and the $F-$stat is extremely large. Unfortunately, running 2SLS produces extremely wide confidence intervals for staffer shares (the 95% CI is roughly 300% wide). This is because we are simultaneously instrumenting for every single staffer's moves, meaning given I staffers, we have $4I$ instruments for I endogenous variables.

We address this issue by adopting and constructing a stronger first stage with machine learning methods. Instead of predicting the first stage using OLS, we instead use an XGBoosted random forest.[3] In Table A.4, we show that this meaningfully improves the predictive power of the instruments. Column 1 shows the root mean square error and accuracy of the linear first stage, both in and out of the sample. Although the accuracy appears high, recall that a staffer can only be in one destination office at any period; thus, a model that always guesses 0 as the outcome would achieve an accuracy of $\frac{1}{436} = 99.77\%$.[4] Thus, the random forest model in Column 2 performs meaningfully better than the linear model. The upside of the random forest is that it has better predictive power and a stronger first stage. The downside is that it is less interpretable: it is difficult to be sure that the instruments are being used by the model, rather than other features provided to it. We provide suggestive evidence that this is not the case in Column 3, where we show that re-estimating the random forest without instruments provides a worse fit both in and out of sample. This suggests that the model is effectively using information provided by the instruments. To formally test for this, we conduct a permutation test. We randomly permute the instruments across observations and re-estimate the random forest model 1000 times. In every single case, out-of-sample performance using the true instruments beats the model using permuted instruments, establishing that the predictive power of

[3] Dikkala et al. (2020) provides a theoretical justification for non-parametric IV and, in particular, finds that the random forest performs best when compared against a number of other non-parametric estimators.

[4] There are 435 Representatives in the House at any given point in time, plus the outside option of not being in Congress.

the random forest first stage is at least partially driven by the instruments.

To complete the IV estimation, we plug in first-stage estimates from the random forest into the design matrix and re-estimate the linear decomposition following the process described above. The one technical note is that the Bayesian bootstrap must additionally account for the first stage. To keep inference conservative, we independently draw bootstrap weights across the first and second stages and take their product.

Appendix B.2 Bayesian estimation

Hamiltonian Monte Carlo We make use of Bayesian methods for inference. Specifically, we utilize Hamiltonian Monte Carlo (HMC) (Gelman et al., 1995), a Markov Chain Monte Carlo (MCMC) method. Traditional Gibbs and Metropolis algorithms exhibit random walk behavior, causing slow progress as they zigzag through the target distribution. While re-parameterization and refined jumping rules can help, this inefficiency persists in complex, high-dimensional models. HMC enhances the Metropolis algorithm by introducing “momentum" variables, enabling iterations to cover more distance in the parameter space and mix more efficiently. Specifically, for each parameter θ_j in the target space, a corresponding “momentum" variable ρ_j is added. Both θ and ρ are updated together using a modified Metropolis algorithm, where the jumping distribution for θ is influenced by ρ. Each HMC iteration involves multiple steps, during which the position and momentum evolve according to rules that mimic physical motion, allowing for the algorithm to traverse the parameter space more quickly, including turning corners to maintain the trajectory's total “energy".

Specifically, HMC augments the posterior density $p(\theta|y)$ with an independent distribution $p(\rho)$ for “momentum" variables, creating a joint distribution $p(\theta, \rho|y) = p(\rho)p(\theta|y)$. We simulate this joint distribution but focus only on the θ values. Again, the momentum variable ρ helps the algorithm move more efficiently through the parameter space. HMC requires the gradient of the log-posterior density, which must be calculated analytically for efficiency. The momentum ρ is usually assigned a multivariate normal distribution with mean 0 and a covariance matrix M that can be conditioned via warm-up. This setup allows the components of ρ to be independent.

An HMC iteration involves three main steps:

1. Updating Momentum: Draw ρ from its prior distribution $N(0, M)$.
2. Leapfrog Steps: Update θ and ρ through a series of “leapfrog" steps:
    - Half-step for ρ: Update ρ using the gradient of the log-posterior.
    $$\rho \leftarrow \rho + \frac{1}{2}\epsilon \frac{d \log p(\theta|y)}{d\theta}$$
    - Full-step for θ: Update θ using the momentum ρ.
    $$\theta \leftarrow \theta + \epsilon M^{-1}\rho$$
    - Another Half-step for ρ: Update ρ again using the gradient.
    $$\rho \leftarrow \rho + \frac{1}{2}\epsilon \frac{d \log p(\theta|y)}{d\theta}$$

3. Accept/Reject Step: Compute the acceptance ratio r based on the joint densities before and after the leapfrog steps. Accept the new θ with probability $\min(r, 1)$; otherwise, retain the previous θ.

This process is repeated until the MCMC algorithm converges, typically assessed by diagnostics like the R-hat statistic and effective sample size.

For challenging HMC problems, it's ideal for tuning parameters to adapt as the algorithm explores the posterior distribution. This involves scaling the mass matrix M to the local curvature of the log density, adjusting the step size ϵ to be smaller in high-curvature areas, and setting the number of steps to be sufficient for effective exploration without excessive circling. Extensions of HMC, like the no-U-turn sampler (NUTS), achieve this adaptive tuning while maintaining detailed balance. In NUTS, the number of steps is determined adaptively during each iteration, continuing until the trajectory reverses direction. This ensures the trajectory explores the parameter space efficiently. NUTS also includes a procedure to adaptively set M and ϵ during a warm-up phase, which is then fixed for the remaining iterations. We use Stan to run HMC with NUTS. Although tuning HMC can be tricky, NUTS simplifies this by eliminating the need to predefine the number of steps.

Our baseline model uses four independent chains, each run for 2000 iterations (with the first 1000 serving as warm-up/burn-in). We typically set δ (the target average proposal acceptance probability) to 0.95,[5] and the seed for sampling to 1. Our baseline prior for both μ_γ, μ_α is N(0,10), and our baseline prior for $\sigma_\gamma, \sigma_\alpha, \sigma_\tau, \sigma$ is N(0,25).

Appendix B.3 Comparison with other approaches and simulation

In this section, we provide details to the discussion of other leading approaches in the economics literature contained in Section 4.

Chan (2021) uses an empirical Bayes procedure to estimate doctor productivity in a setting where junior and senior physicians are paired together in teams of two. He uses a discrete type space, allowing for 6 types of juniors and 9 types of seniors (based on tenure), and estimates a separate variance for each pair of types. Extending this approach to our setting would require estimating a separate variance for each feasible combination of types in a team (6 types of staffers among the 18 possible "juniors" in a team, and 9 types for the senior, the member of Congress). Assuming symmetry across individuals in a team requires the estimation of $9 * \binom{24}{6} = 1.2$ million parameters; without symmetry, this is $6^{18} * 9 = 1$ quadrillion parameters.

Bonhomme (2021) estimates researcher productivity using both a linear and non-linear model. For his nonlinear model, he uses a variational Bayes (ELBO) procedure and restricts attention to teams no larger than two. He estimates his model with K types of workers (using $K = 4$ at baseline). Evaluating the ELBO objective function requires a computing a nested sum containing (1) the number of individuals, (2) the number of teams, and (3) the number of types for each individual on the team. For the simplest case of a single office containing 18 staffers, this is $18 * 4^{18} = 1$ trillion operations. This

[5] This is higher than the default used in packages such as RStan, as it produces results that are more robust when posteriors have high curvature, at the cost of computational speed.

is a lower bound as there are many unique teams in the data (on the order of 10,000), although not all of them hire the maximal 18 staffers. Actually estimation typically requires many calls to the objective function.

Ahmadpoor and Jones (2019) pose a nonlinear least squares problem that they solve using simple gradient descent. In particular, they pick an informed initialization, run gradient descent for one hundred thousand iterations, and conclude the procedure. When implementing their procedure to our nested CES problem, we made use of the same learning rates as the original authors, but also experimented with a range of different rates. We were unable to find a clearly better set of hyperparameters, and found results to be highly sensitive to initial parameterization. In particular, we used the same history of moves as our primary sample, and simulated staffer and representative effects using the same distributions as the baseline estimates presented in Table 3, Panel B, with a small amount of noise (a standard normal draw). We then initialized all parameters to the CES problem at the oracle solution (in particular, with $\rho_{\text{staff}} = 0.794$). The method returns a result that is substantially different: $\hat{\rho}_{\text{staff}} = 9.461$, and the average value of α compared to $\hat{\alpha}$ is off by roughly two orders of magnitude. Finally, we note that this procedure is also slower than the Bayesian one that we adopt. Although a single run of their estimation procedure to recover a point estimate takes roughly twice as long as ours, inference via 100 iterations of the Bayesian bootstrap would increase total runtime by a factor of 100.

As discussed in the main text, Bergeron et al. (2022) adopts a linear model similar to ours. However, because output is non-parametrically estimated as a combination of discrete types, this would still result in over 200,000 parameters to estimate in our context, above the number of observations that we have.

In addition to these papers, we are also aware of Freund (2024), Iranzo et al. (2008), and Weidmann et al. (2024), who all study the contribution of individuals in team environments. Each of these papers directly estimates an individual's effect, and uses these estimates to study their impact on teams. Iranzo et al. (2008) estimate individual effects through a standard AKM wage regression, and plug these effects into their firm production estimation procedure. Freund (2024) estimates individual effects through an augmented AKM regression that takes the team into account, but the outcome data are ultimately still individual-level wages. Weidmann et al. (2024) directly measures individual effects in the lab before randomly pairing individuals into teams.

Finally, we discuss the simulation procedure that we use to validate our estimates. For the dynamic models, we simulate staffer and Representative effects using the same distributions as the baseline estimates presented in Table 3, Panel A. We then construct a panel of movers that matches the main mover sample along the following moments: the average and standard deviation for number of quarters spent in an office, the average and standard deviation for number of moves between offices, and the average number of quarters that a Representative spends in Congress. This leads to a Representative share of differences between above and below median offices of 73%, and a Representative share of differences between destination and origin offices of 8 to 12% (depending on whether odd or even quarters are used). In estimation, we find a Representative share of differences between above and below median offices of 72%, with a 95% CI of [0.713, 0.733]. We plot the dynamic estimate of the Representative share of differences between destination and origin offices in Figure A.12, Panels A and B, and find them consistent with the true

values. We also plot predicted staffer, Representative, and other team-member moves in Figure A.12, Panels C and D, and find results consistent with the discussion of Equation 6. For our Bayesian procedure, we note that, by design, inference on the posterior is correct. Nevertheless, we evaluate the nested CES estimation procedure and find that the true parameter values fall within the 95% credible interval.

Appendix C Tables and Figures

Table A.1: Correlating staffer FEs with each other

FE source:	Bills cospons.	Exit comm.	Pass House	Laws passed
	(1)	(2)	(3)	(4)
OLS: # orig. bills cospons.	1.000			
	(.)			
OLS: # bills exit committee	0.358***	1.000***		
	(0.039)	(0.000)		
OLS: # bills pass House vote	0.528***	0.570***	1.000***	
	(0.037)	(0.040)	(0.000)	
OLS: # bills becoming law	0.537***	0.592***	0.985***	1.000***
	(0.038)	(0.040)	(0.007)	(0.000)

Notes: Regressions are at the staffer level. FEs are sourced from an OLS estimate of Equation 1. Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.2: Misspecified AKM designs, staffer share of differences

Outcome:	Bills cospons.	Laws passed	Exit comm.	Pass House	Important bills
	(1)	(2)	(3)	(4)	(5)
Panel A: baseline, team-based AKM model					
Staffer share	0.489	0.240	0.212	0.395	0.141
	[0.416,0.591]	[0.055,0.458]	[0.028,0.416]	[0.152,0.575]	[-0.143,0.357]
Panel B: misspecified, standard AKM model (all staffers)					
Staffer share	0.233	0.034	0.022	-0.019	0.066
	[0.157,0.302]	[-0.083,0.130]	[-0.090,0.150]	[-0.143,0.116]	[-0.051,0.176]
Panel C: misspecified, standard AKM model (only Chiefs of Staff)					
Staffer share	0.215	0.055	0.080	0.047	0.072
	[0.188,0.248]	[0.020,0.145]	[0.010,0.167]	[-0.036,0.104]	[-0.014,0.154]
Panel D: misspecified, standard AKM model (only Legislative Directors)					
Staffer share	0.218	0.049	0.050	0.047	0.006
	[0.171,0.165]	[-0.067,0.099]	[-0.005,0.109]	[-0.017,0.197]	[-0.140,0.161]

Notes: This table presents the staffer share of differences between above and below-median offices across a number of legislative outcomes. In Panel A, we present the baseline estimates from the team-based AKM model as in Figure 2. Panels B-D present misspecified models that run the standard AKM design at the individual level. Panel B includes all movers, meaning an observation is repeated as many times as there are staffers in the office. Panels C-D restrict there to be a single staffer per office. In Panel C, this is the Chief of Staff. In Panel D, this is the Legislative Director. Standard errors are computed via 100 iterations of the Bayesian bootstrap. 95% confidence intervals are displayed beneath point estimates in brackets.

Table A.3: IV — linear first stage

	(1)	(2)	(3)	(4)
Stay in same office	0.807***	0.807***	0.798***	0.798***
	(0.002)	(0.002)	(0.002)	(0.002)
Turnover	0.021***	0.021***	0.021***	0.021***
	(0.005)	(0.005)	(0.005)	(0.005)
Total moves from orig. to dest.	0.019***	0.018***	0.019***	0.019***
	(0.001)	(0.001)	(0.001)	(0.001)
# leaving dest.	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Moves from O-D X # leave dest.	0.006***	0.006***	0.006***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Dist. in party share	0.000	0.000	0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
N	21018181	21018180	21018180	21018180
F	42942.09	43095.46	36766.30	37176.89
RMSE	0.031	0.031	0.031	0.031
FE	None	Time	+ Orig./Dest.	+ Staffer

Notes: This table presents regression coefficients and standard errors for the first stage of the IV at the staffer-origin office-destination office-quarter level. Column 2 adds time-fixed effects, column 3 adds origin and destination-fixed effects, and column 4 adds staffer-fixed effects. Standard errors are clustered at the staffer-quarter level. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.4: IV — comparing first stages

	OLS	RF	RF (no instruments)	(2) < (3) $p-$val.
	(1)	(2)	(3)	(4)
RMSE (in sample)	0.0310	0.0273	0.0274	-
Accuracy (in sample)	99.894%	99.915%	99.914%	-
RMSE (out of sample)	0.0312	0.0266	0.0280	< 0.001
Accuracy (out of sample)	99.892%	99.920%	99.912%	> 0.999

Notes: This table presents in-sample and out-of-sample root mean square error (RMSE) and accuracy for various IV first stages. Column 1 presents results from OLS (Table A.3, Column 4). Column 2 presents results from an XG-Boosted random forest. Column 3 presents results from the random forest without instruments. $p-$values in column 4 are computed using a permutation test (permuting instrument rows then re-fitting the random forest) with 1000 draws.

Table A.5: Bayesian model, HMC NUTS sampler properties

Property	Min.	25%	Median	Mean	75%	Max
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: baseline model						
Accept stat.	0.00	0.84	0.94	0.85	0.98	1.00
Step size	0	0.001	0.001	0.002	0.002	14.
Tree depth	0.0	8.0	8.0	7.9	8.0	10.0
N leapfrog	1	255	255	296	255	1023
Divergent	0.000	0.000	0.000	0.052	0.000	1.000
Panel B: CES model						
Accept stat.	0.00	0.94	0.98	0.94	1.00	1.00
Step size	0	0.0015	0.0079	0.0126	0.0089	10.95
Tree depth	0.0	9.0	9.0	8.9	10.0	10.0
N leapfrog	1	511	511	662	1023	1023
Divergent	0.000	0.000	0.000	0.016	0.000	1.000

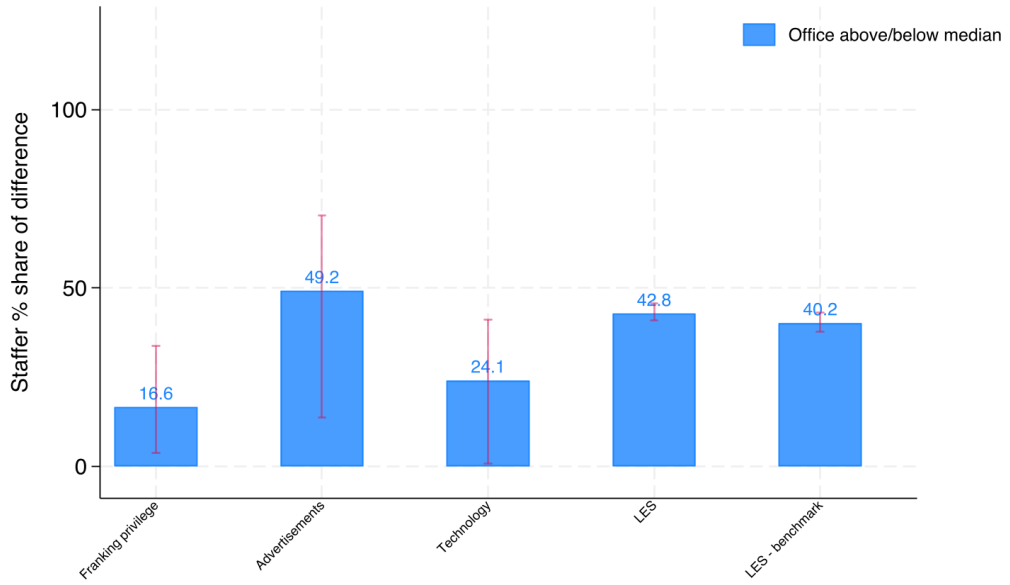
Notes: This table presents sampler properties from the baseline model (Panel A) and CES model (Panel B). Columns 2 through 7 display the minimum, 25th percentile, median, mean, 75th percentile, and maximum respectively. Accept stat. is the probability that a proposal is accepted. Step size controls how far the momentum variable updates. Tree depth is the depth of the binary tree used by NUTS, which determines the number of leapfrog steps taken by the sampler to avoid random walk behavior. Divergent transitions are instances when the Hamiltonian value deviates substantially from the initial value, indicating potential random walk behavior.

Table A.6: Correlation between Bayesian ideology estimates and DW-Nominate scores

	1st Dimension		2nd Dimension	
	Dem.	Rep.	Dem.	Rep.
	(1)	(2)	(3)	(4)
DW-Nominate score	-0.0144	0.0184**	0.1468*	0.0296**
	(0.0756)	(0.0089)	(0.0845)	(0.0126)

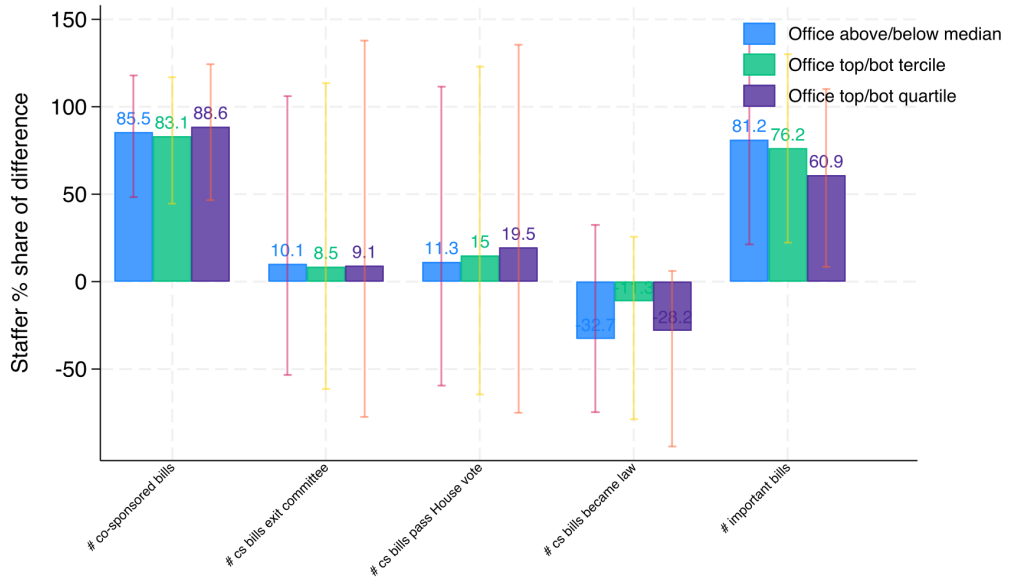
Notes: This table presents regressions at the Representative level, where the outcome is the posterior mean of a Representative's ideology when writing bills, and the independent variable is the Representative's DW-Nominate score from roll call votes. Columns 1 and 2 present results from the 1st dimension of DW-Nominate, and columns 3 and 4 from the 2nd dimension. Odd rows restrict the sample to Democrats, even rows to Republicans. Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

Figure A.1: Staffer share of differences in office spending, linear decomposition



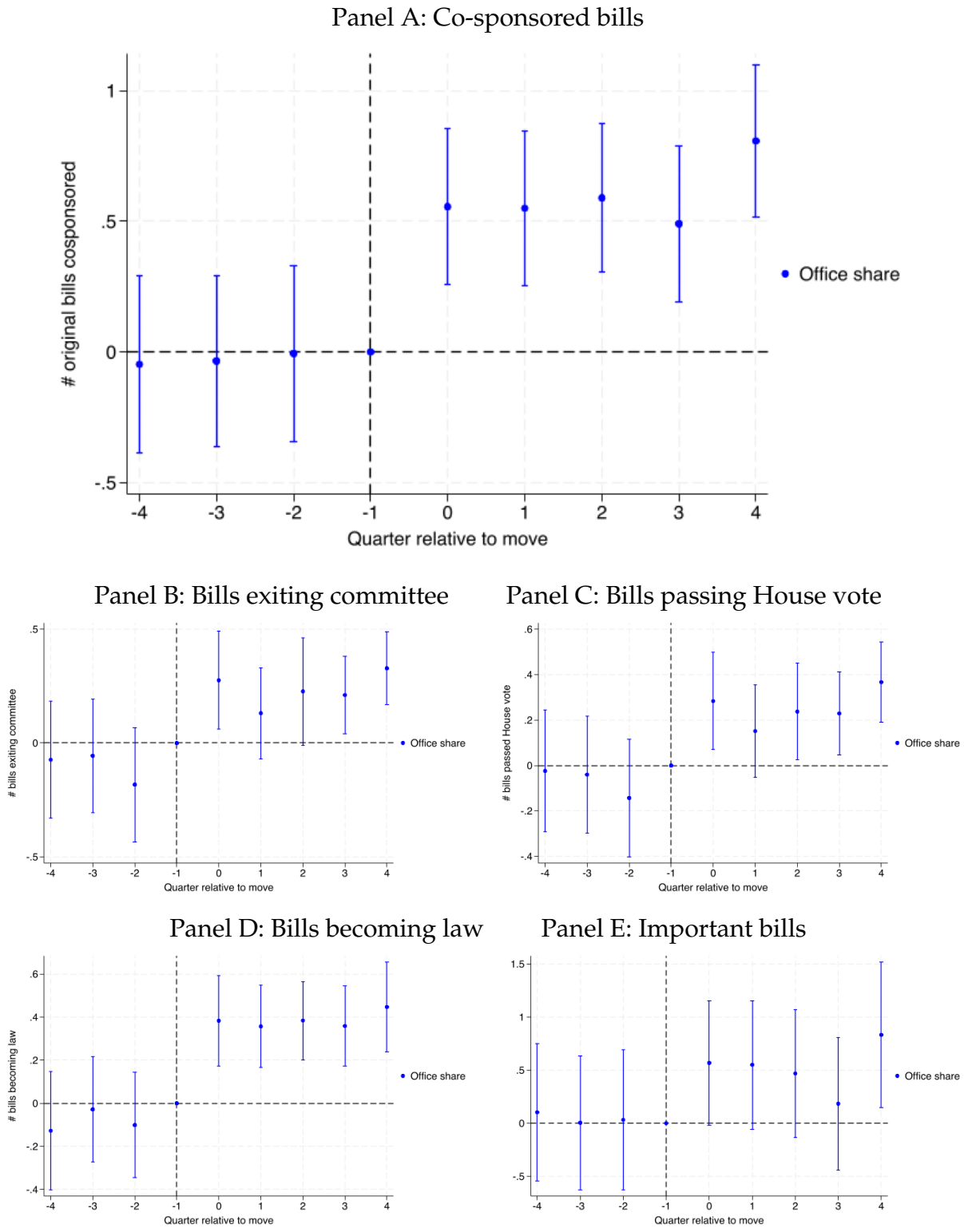
Note: This figure plots the staffer share of differences in office spending (first three columns) and legislative productivity (last two columns), as well as 95% confidence intervals. The staffer share is computed from fixed effects estimated in a team-based mover design framework (Equation 1). LES refers to the Legislative Effectiveness Score.

Figure A.2: Staffer share of differences in office spending, linear decomposition (IV)



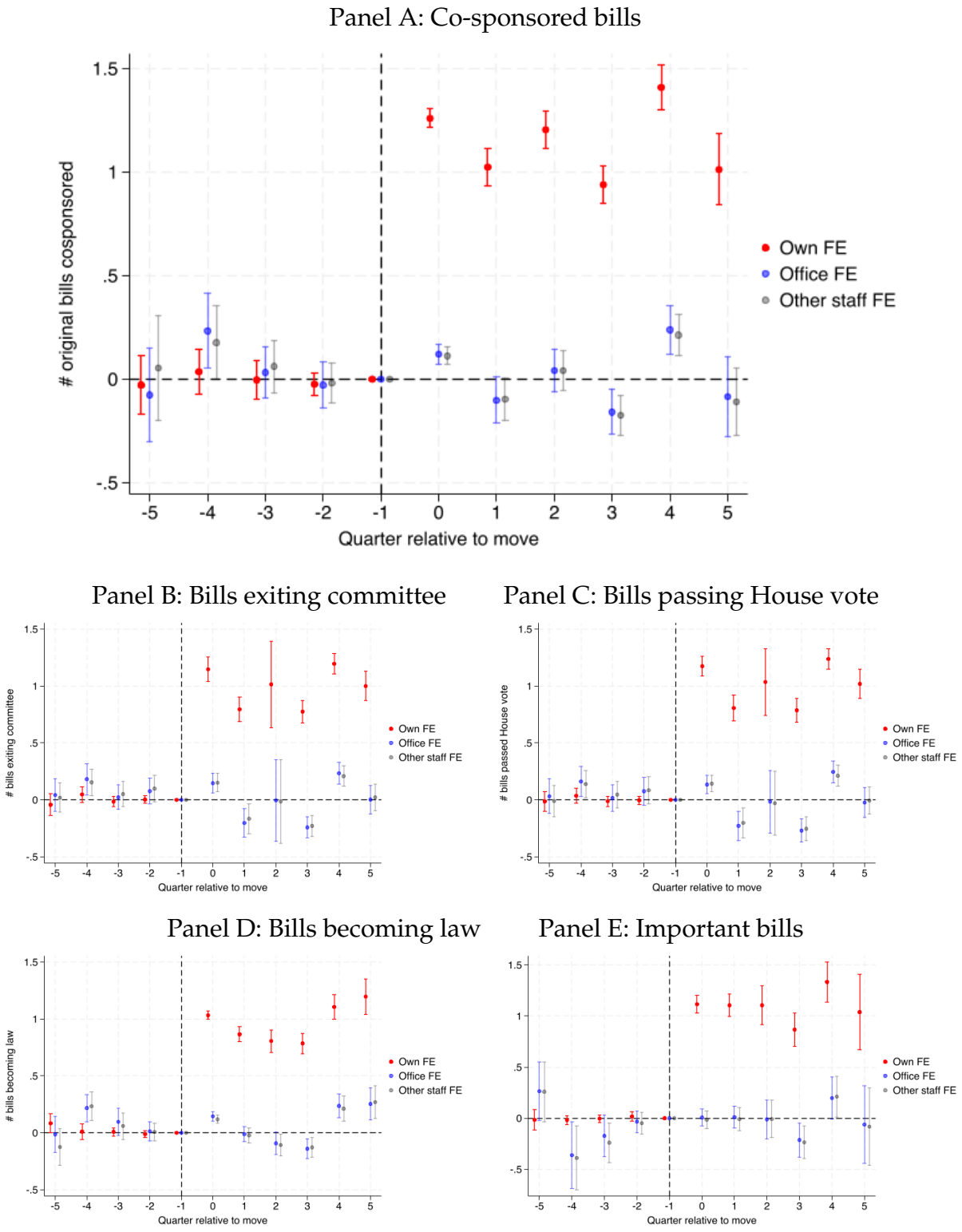
Note: This figure plots the staffer share of differences between offices that are above and below median in legislative output in blue, as well as 95% confidence intervals. The staffer share is computed from instrumented fixed effects estimated in a team-based mover design framework (Equation 1), using a random forest first stage. Staffer shares for the difference in top and bottom tercile (quartile) offices are plotted in green (purple). Bars are arranged from left to right in order of closeness to passage into law, except for 'important bills', which refers to the number of bills referenced in the Congressional Quarterly Almanac year-end write-up. 'cs bills' refers to the number of bills co-sponsored by the office.

Figure A.3: Dynamic estimation of staffer shares, in-sample (odd quarters)



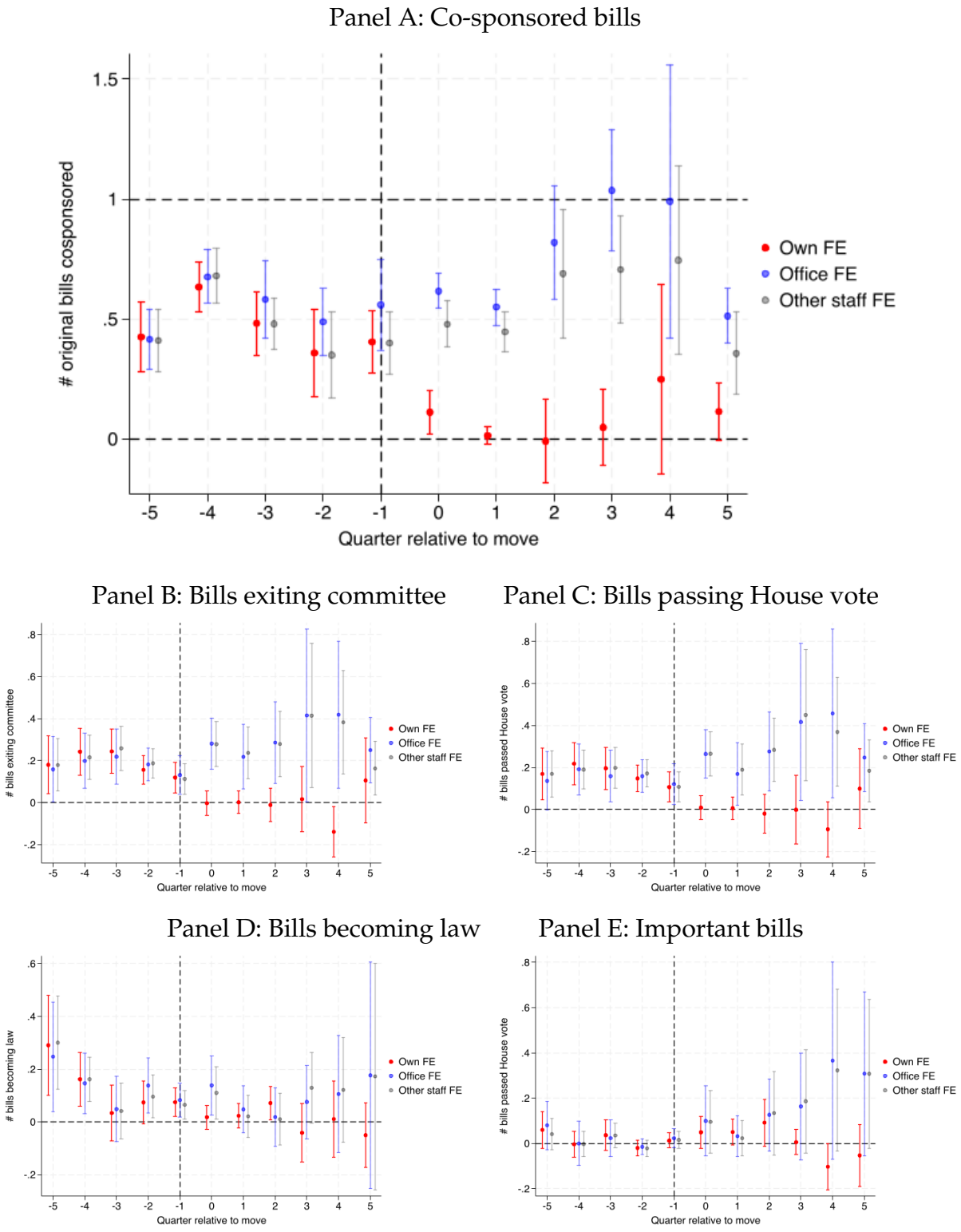
Note: This figure plots coefficients and 95% confidence intervals for an event study specification (Equation 5) at the staffer-quarter level (for odd quarters only). The x-axis shows quarters to a staffer's first move. Displayed coefficients are the Representative share of differences between the destination and origin office's legislative output. The specific type of legislative output varies by panel. Standard errors are clustered at the Representative level.

Figure A.4: Dynamic estimation of staffer movement, in-sample (odd quarters)



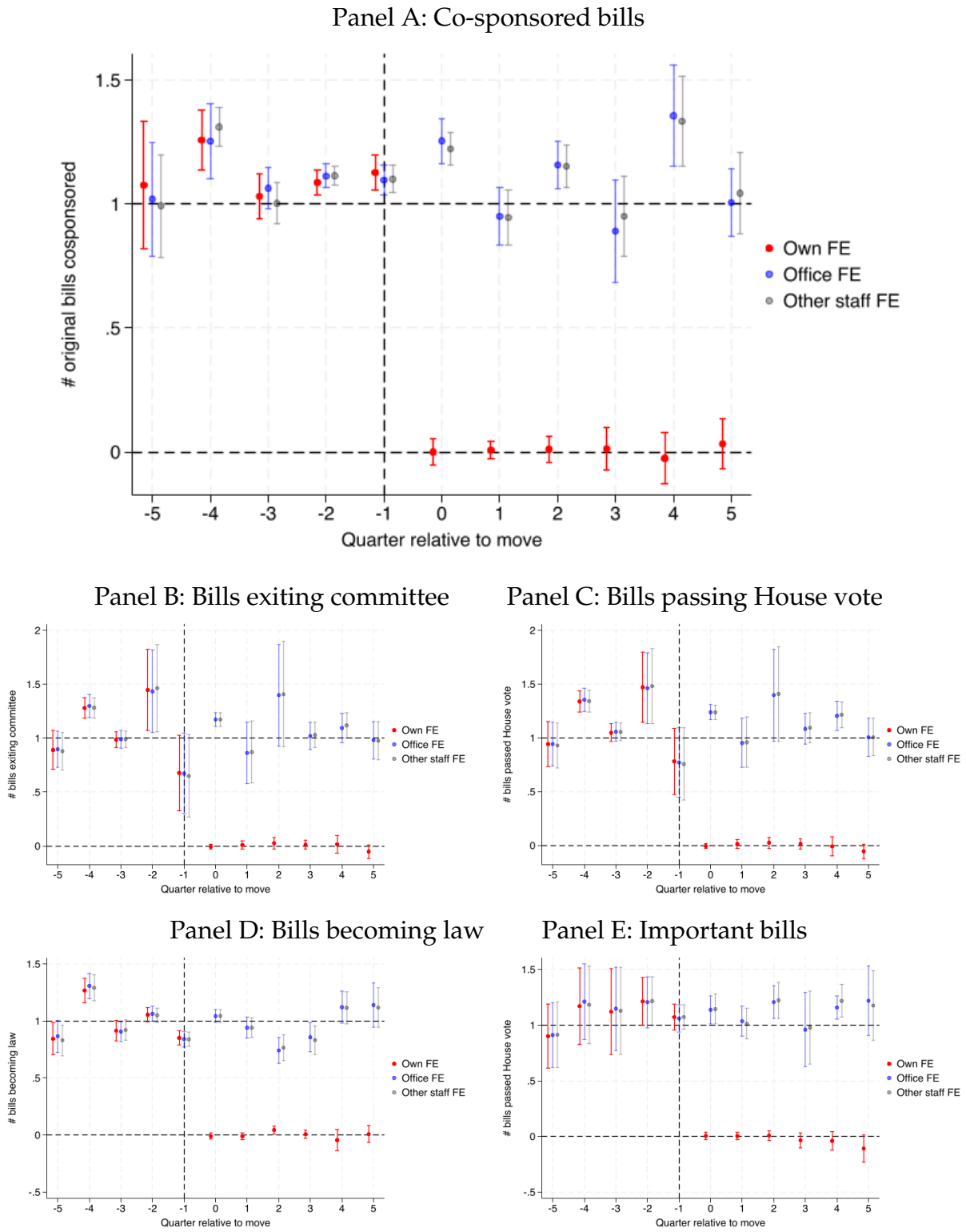
Note: This figure plots coefficients and 95% confidence intervals for an event study specification (Equation 6) at the staffer-quarter level (for odd quarters only). The x-axis shows quarters to a staffer's first move. The outcome is a specific type of destination office legislative output that varies by panel. Coefficients on the moving staffer's fixed effect are plotted in red, coefficients on the destination Representative's fixed effect are plotted in blue, and coefficients on the destination team of staffers' fixed effect are plotted in gray. Standard errors are clustered at the Representative level.

Figure A.5: Dynamic estimation of staffer movement, from origin office



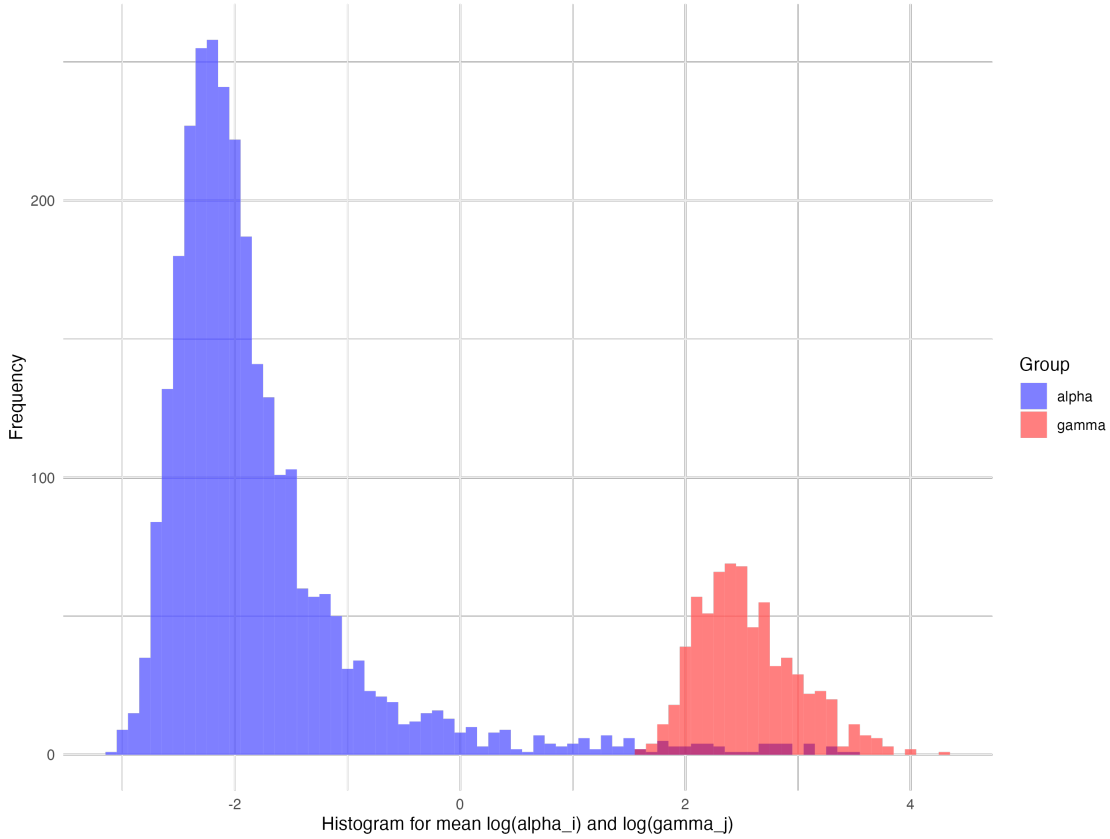
Note: This figure plots coefficients and 95% confidence intervals for an event study specification (Equation 6) at the staffer-quarter level (for even quarters only). The x-axis shows quarters to a staffer's first move. The outcome is a specific type of origin office legislative output that varies by panel. Coefficients on the moving staffer's fixed effect are plotted in red, coefficients on the destination Representative's fixed effect are plotted in blue, and coefficients on the destination team of staffers' fixed effect are plotted in gray. These fixed effects are estimated out of sample on odd quarters. Standard errors are clustered at the Representative level.

Figure A.6: Dynamic estimation of staffer movement, from origin office, in-sample (odd quarters)



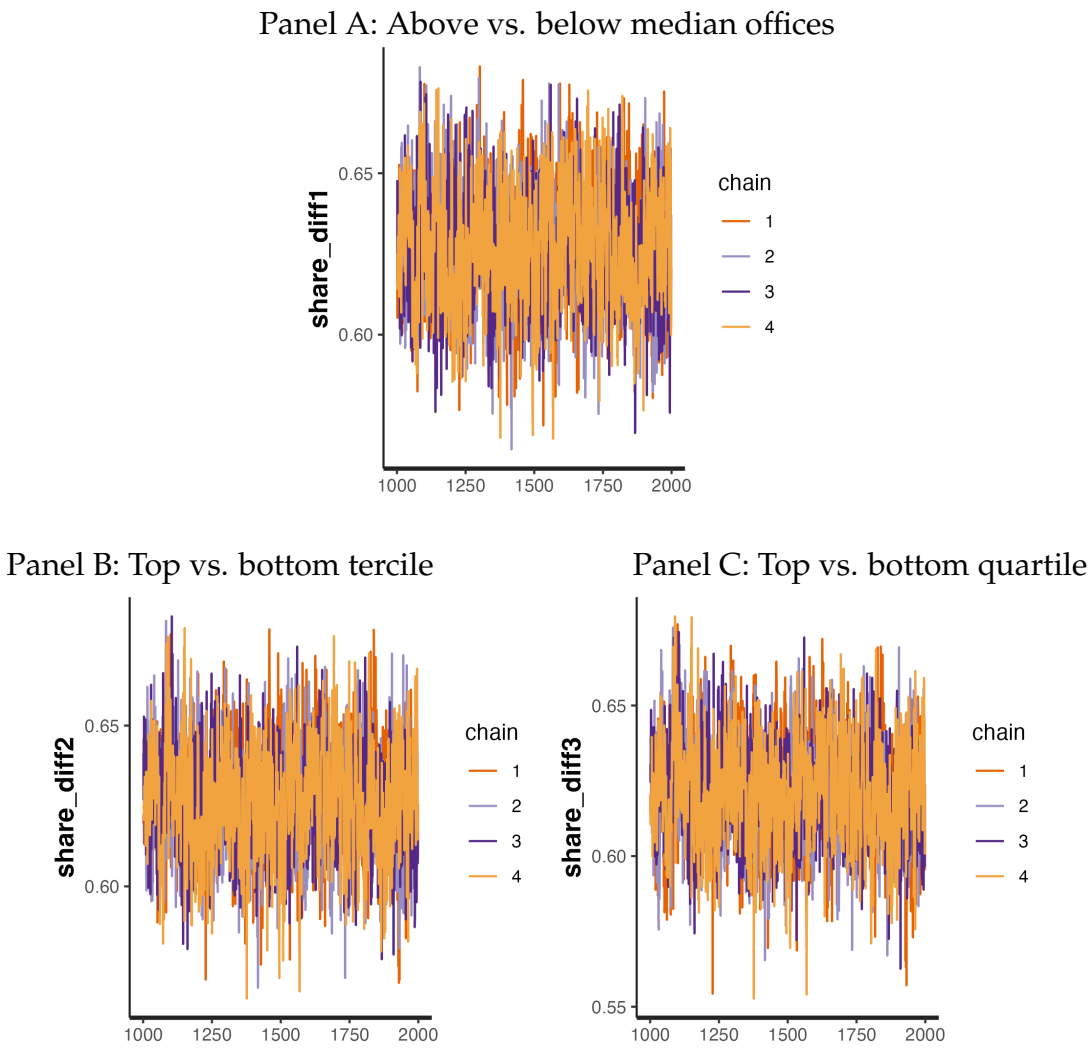
Note: This figure plots coefficients and 95% confidence intervals for an event study specification (Equation 6) at the staffer-quarter level (for odd quarters only). The x-axis shows quarters to a staffer's first move. The outcome is a specific type of origin office legislative output that varies by panel. Coefficients on the moving staffer's fixed effect are plotted in red, coefficients on the destination Representative's fixed effect are plotted in blue, and coefficients on the destination team of staffers' fixed effect are plotted in gray. Standard errors are clustered at the Representative level.

Figure A.7: Posterior distributions of α_i, γ_j from Bayesian estimation



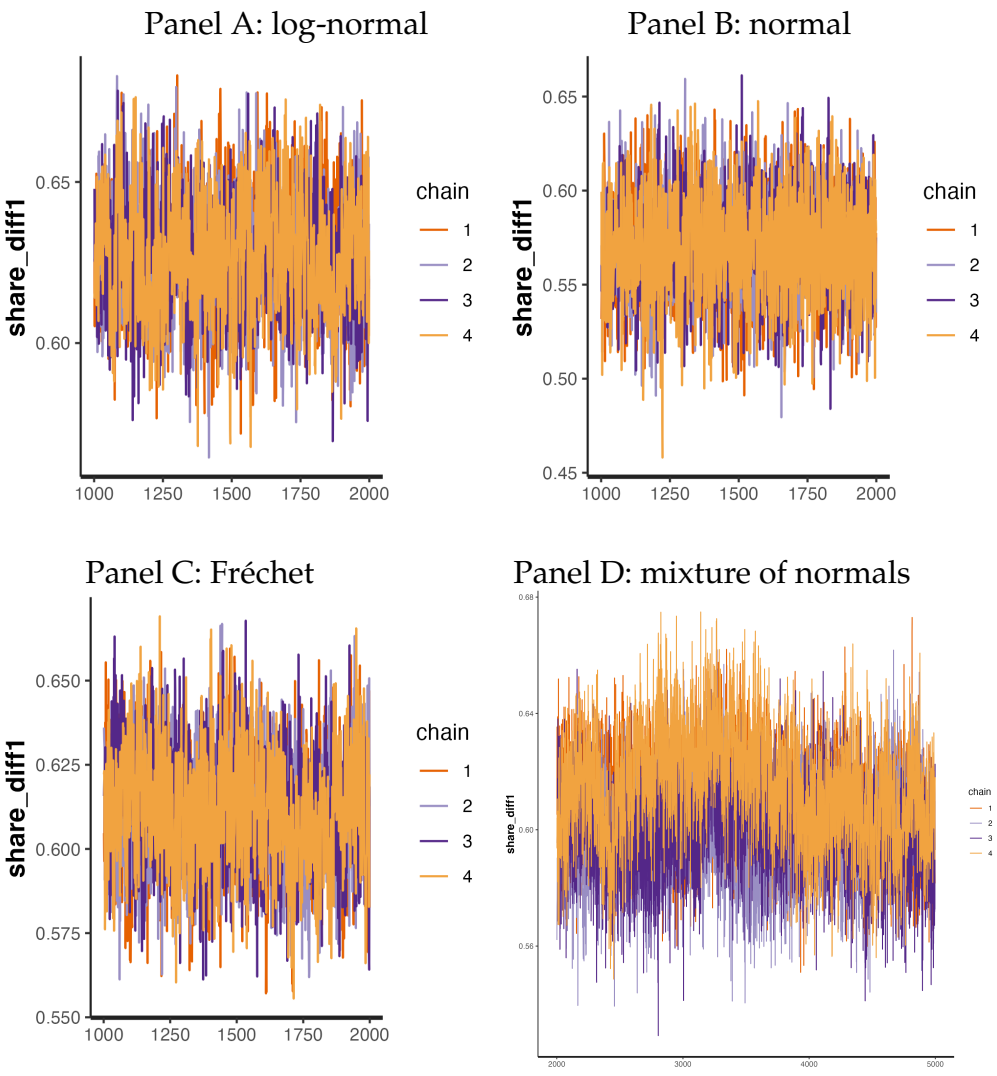
Note: This figure plots the posterior mean of $\log(\alpha_i)$ (in blue) and $\log(\gamma_j)$ (in red) for each staffer and Representative. Estimates for α_i and γ_j come from a Bayesian estimation of Equation 1.

Figure A.8: Trace plots for Representative share of difference in number of bills co-sponsored, varying share



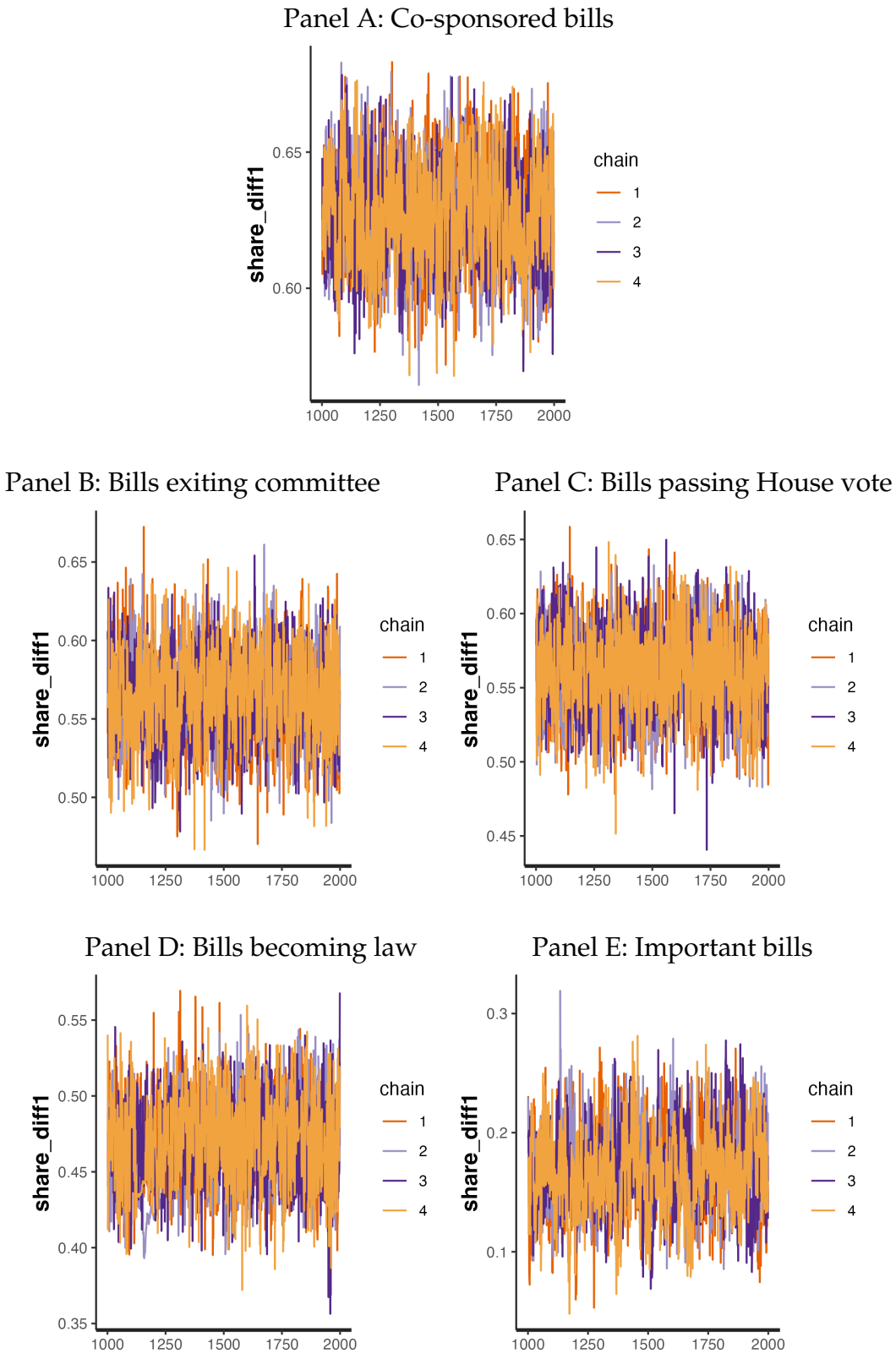
Note: This figure presents trace plots for a Bayesian estimation of the team-based mover design framework (Equation 1). The x-axis represents samples from a posterior drawn via MCMC. Panel A shows the Representative share of differences between above and below-median offices in bills produced. Panels B and C display the Representative share of differences between above and below tercile and quartile offices in bills produced. Values for 4 different chains are displayed.

Figure A.9: Trace plots for Representative share of difference in number of bills co-sponsored, varying distribution



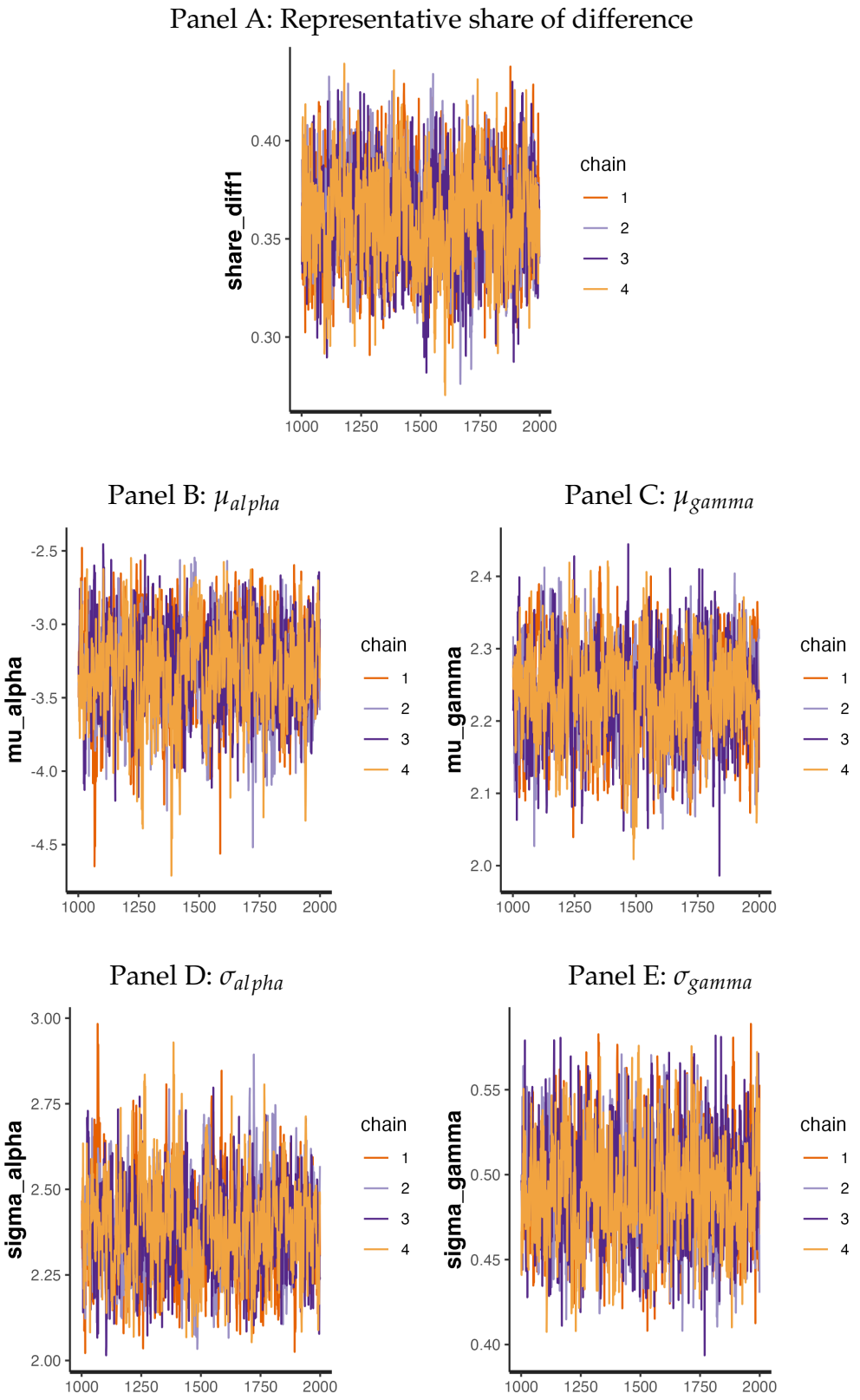
Note: This figure presents trace plots for a Bayesian estimation of the team-based mover design framework (Equation 1). The x-axis represents samples from a posterior drawn via MCMC. Panel A shows the Representative share of differences between above and below-median offices in bills produced when parameterizing staffer and Representative effects using a log-normal distribution. Panels B, C, and D instead use a normal, Fréchet, and mixture of normals. Values for 4 different chains are displayed.

Figure A.10: Trace plots for Representative share of difference, varying outcome



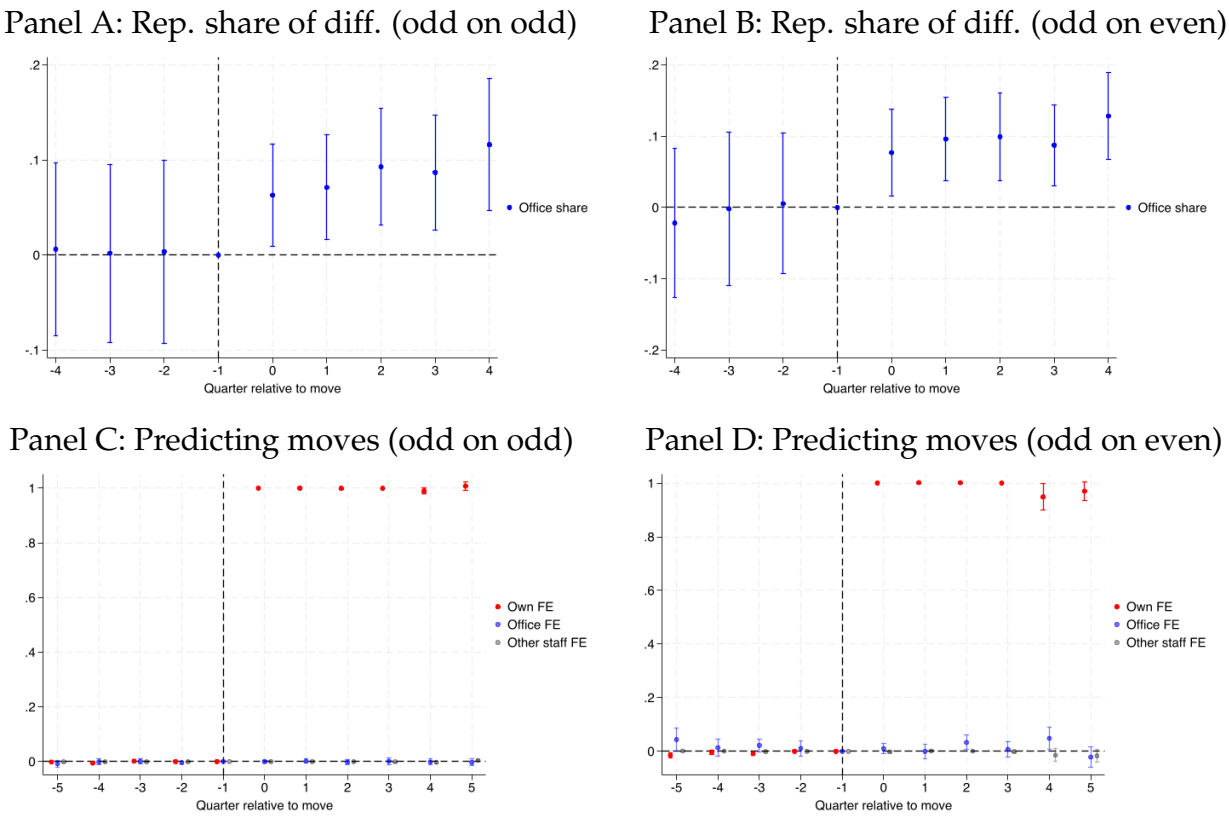
Note: This figure presents trace plots for a Bayesian estimation of the team-based mover design framework (Equation 1). The x-axis represents samples from a posterior drawn via MCMC. Panel A shows the Representative share of differences between above and below-median offices in bills produced. Panels B, C, D, and E instead show the Representative share of differences for the number of co-sponsored bills exiting committees, passing a House vote, becoming law, or being important bills. Values for 4 different chains are displayed.

Figure A.11: Trace plots for the model including non-movers, number of bills co-sponsored



Note: This figure presents trace plots for a Bayesian estimation of the team-based mover design framework (Equation 1). The x-axis represents samples from a posterior drawn via MCMC. Panel A shows the Representative share of differences between above and below-median offices in bills produced. Panels B and D (C and E) display the staffer (Representative) log-normal distribution's mean and standard deviation. All staffers (not just movers) are included in the estimation. Values for 4 different chains are displayed.

Figure A.12: Validation of dynamic estimation by simulation



Note: This figure plots coefficients and 95% confidence intervals for two event study specifications (Equation 5 in Panels A and B, Equation 6 in Panels C and D) at the staffer-quarter level. The underlying data are simulated. The x-axis shows quarters to a staffer's first move. The outcome is bills produced. In Panels A and B, coefficients on the Representative share are plotted. The true values are 8 and 12% respectively. In Panels C and D, coefficients on the moving staffer's fixed effect are plotted in red, coefficients on the destination Representative's fixed effect are plotted in blue, and coefficients on the destination team of staffers' fixed effect are plotted in gray. Standard errors are clustered at the Representative level.