Online Appendix to

"Did U.S. Politicians Expect the China Shock?"

Matilde Bombardini, Bingjing Li, Francesco Trebbi

- Not for Publication -

Empirical Model: Additional Details

A1. Microfoundation of Equation (1)

The expression used in equation (1) is a common representation of expected electoral support (Bartels, 1993; Canen, Jackson and Trebbi, 2019). Here we offer a microfoundation for the expression $h_t(\cdot)$ in the text.

Consider a congressional district with N voters, with $j = 1, ..., N$. The district has candidates i, i' running for election at time $t+1$. We will assume that i is the incumbent candidate in office at t without loss of generality. Excluding an open race and imposing the presence of an incumbent representative is necessary as we explore the electoral consequences of supporting a bill at time t. $d_{i,t}$ defines i's vote decision and it will be based on the expected electoral consequences that this will have at $t+1$.

Voters are assumed to vote based on life events. During normal times in their life, voters employ a random utility framework and evaluate candidates based on each politician's valence and policy position, with voters choosing the candidate that delivers the higher expected utility. In exceptional times, when voters are hit by particularly strong shocks, voters become single-issue voters (Egorov, 2015). Single-issue voters punish or reward the incumbent only based on his/her past vote on the issue of relevance. As we focus on the China shock, adverse local labor market consequences of the China shock are the stochastic event that triggers the switch to become a single-issue voter. This representation of single-issue voters is stark for the sake of simplicity and Cruz et al. (2018) show how state-dependent voter preferences of a more general form can be modeled and estimated.

Define $S_{i,t}$ the proxy for the degree of exposure of the local labor market in the district represented by i at time t to increasing imports from China. Let $f(S_{i,t})$, with function $f(\cdot)$ continuous and increasing, indicate the probability that a voter j faces exceptional times due to the China shock (e.g., j loses her job due to outsourcing; j has a son who cannot find employment in the area, etc.).

With probability $1-f(S_{i,t})$ (i.e., in normal times), a voter j has preferences over the choice of candidates i and i'. Candidates differ in terms of their valence, λ, that is the quality of the candidate, and policy position $\theta \in \Pi$, in a policy/ideology space. Elected politicians advance the policy point θ (Ansolabehere, Snyder Jr and Stewart III, 2001; Lee, Moretti and Butler, 2004). Voters are heterogeneous in their ideal policies with bliss points $q_j \in \Pi$. Both λ and θ for each candidate are known to voters.

Utility of voter j of type q_j to vote for politician i with valence λ and policy stance θ is:

$$u(\lambda, \theta; q_j) = U_i^j(q_j) + \varepsilon_{i,j}$$

where $\varepsilon_{i,j}$ is the random utility component specific to match (i, j). As an example, possible specification for U is $U_i^j(q_j) = \gamma\lambda - |q_j - \theta|^\varsigma - \chi * (\lambda * |q_j - \theta|^\varsigma)$ estimated in Kendall, Nannicini and Trebbi (2015).

The probability voter j votes for i is then:

$$p_{i,j} = \Pr\left[U_i^j\left(q_j\right) + \varepsilon_{i,j} \geq U_{i'}^j\left(q_j\right) + \varepsilon_{i',j}\right]$$

and if we assume extreme value distribution for $\varepsilon_{i,j}$ i.i.d. $F(\varepsilon_{ij}) = exp\left(-e^{-\varepsilon_{ij}}\right)$, then:

$$p_{i,j} = \frac{e^{U_i^j(q_j)}}{\sum_l e^{U_l^j(q_j)}}.$$

With probability $f(S_{i,t})$ (i.e., in exceptional times), voter j rewards politician i if $\mathbb{1}\left\{d_{i,t-1}= \text{vote for } q_{t-1}\right\}$, and votes for i' if instead $\mathbb{1}\left\{d_{i,t-1}= \text{vote for } x_{t-1}\right\}$, so that:

$$p_{i,j} = 1 \times \mathbb{1}\left\{d_{i,t-1}= \text{vote for } q_{t-1}\right\}.$$

The expected electoral support for politician i can now be calculated. Upon the realization of $S_{i,t+1}$, each voter j has a different probability of voting in favor of politician i, $p_{i,j}$, and the vote choice is a non-identically distributed independent Bernoulli random variable:

$$v_{i,j,t+1} = \begin{cases} 1 & with\ p_{i,j} \\ 0 & with\ 1 - p_{i,j} \end{cases}$$

The sum of votes in support of i in the district is the random variable $V_{i,t+1} = \sum_{j=1}^{N} v_{i,j,t+1}$. $V_{i,t+1}$ is distributed as a Poisson Binomial distribution, as it is the convolution of non-i.i.d. Bernoulli random variables. The Poisson Binomial is governed by the parameter $P_{i,t+1} = \sum_j p_{i,j}$. By Le Cam (1960)'s Theorem, the Poisson Binomial distribution is bound by the Poisson distribution with parameter $P_{i,t+1}$. This implies that the expected number of votes for candidate i conditional on $S_{i,t+1}$ can be approximated by $P_{i,t+1}$.

Using the expressions above for period $t+1$:

$$\begin{aligned}
P_{i,t+1} &= \sum_j p_{i,j} \\
&= \sum_j \left[\frac{e^{U_i^j(q_j)}}{\sum_l e^{U_l^j(q_j)}} \left(1 - f(S_{i,t+1})\right) + \mathbb{1}\left\{d_{i,t} = \text{vote for } q_t\right\} f(S_{i,t+1}) \right] \\
&= \sum_j \frac{e^{U_i^j(q_j)}}{\sum_l e^{U_l^j(q_j)}} - \sum_j \frac{e^{U_i^j(q_j)}}{\sum_l e^{U_l^j(q_j)}} f(S_{i,t+1}) \times \mathbb{1}\left\{d_{i,t} = \text{vote for } x_t\right\} \\
&+ \sum_j \left(1 - \frac{e^{U_i^j(q_j)}}{\sum_l e^{U_l^j(q_j)}}\right) f(S_{i,t+1}) \times \mathbb{1}\left\{d_{i,t} = \text{vote for } q_t\right\} \\
&\simeq \gamma^0 + \gamma^1 S_{i,t+1} \times \mathbb{1}\left\{d_{i,t} = \text{vote for } x_t\right\} + \gamma^2 S_{i,t+1} \times \mathbb{1}\left\{d_{i,t} = \text{vote for } q_t\right\} \\
&= \gamma^0 + \gamma^2 S_{i,t+1} + \left(\gamma^1 - \gamma^2\right) S_{i,t+1} \times \mathbb{1}\left\{d_{i,t} = \text{vote for } x_t\right\},
\end{aligned}$$

where the third step comes from the linear approximation of $f(S_{i,t+1})$, with $\gamma^0 \simeq \sum_j \frac{e^{U_i^j(q_j)}}{\sum_l e^{U_l^j(q_j)}}$, $\gamma^1 \simeq -\sum_j \frac{e^{U_i^j(q_j)}}{\sum_l e^{U_l^j(q_j)}}$ and $\gamma^2 \simeq \sum_j \left(1 - \frac{e^{U_i^j(q_j)}}{\sum_l e^{U_l^j(q_j)}}\right)$. The last step is the function that we employ in Equation (1) in the text, and in light of the model $\gamma_1 - \gamma_2 < 0$.

Some evidence in support of the microfoundation

To provide explorative evidence for the electoral channels proposed here, we relate the electoral outcomes of the incumbents to their voting records on the NTR (and PNTR) bills over the time frame of our analysis.[41] The following equation aims at estimating the differential electoral losses associated to supporting China's NTR status for politicians representing districts adversely impacted by the China shock:

$$ShareVote_{i,t} = \beta_1 S_{i,t,t-2} \times VoteProCHN_{i,t,t-2} + \beta_2 S_{i,t,t-2} + \beta_3 VoteProCHN_{i,t,t-2} + \mathbb{X}'_{i,t}\beta_3 + \gamma_{s,t} + u_{i,t},$$

where $ShareVote_{i,t}$ is the share of votes obtained by incumbent i in the election in year t. $VoteProCHN_{i,t,t-2}$ is an indicator equals to 1 if the incumbent only voted in favor of China during the congressional session (i.e., over the period $t-2$ and t). $S_{i,t,t-2}$ is the realized China Shock over $t-2$ and t. The vector $\mathbb{X}_{i,t}$ contains the individual and district characteristics, including DW Nominate score, tenure and party affiliation of the politician, and manufacturing employment share in the district and its interaction with the voting record. $\gamma_{s,t}$ denotes the state-year fixed effects. β_1 qualitatively corresponds to $\gamma_1 - \gamma_2$ in our theoretical framework,

[41]Data on votes received by different candidates are obtained from MIT Election Data and Science Lab (2017).

which is hypothesized to be negative. The standard errors are two-way clustered at the state and individual politician level.

The regression results are reported in Table A.1. As shown in column (1), voting in favor of China reduces the incumbent's vote share, and more so when the district is more exposed to the import shock from China in the past 2 years. Column (2) further controls the individual politician fixed effects, the estimated coefficient for β_1 remains negative but becomes imprecisely estimated. In columns (3) and (4), we repeat the analysis but replace the 2-year import shock by the 5-year import shock, and the voting records in the past 2 years by the voting records in the past 5 years. The estimated coefficients for the interaction term are negatively significant across specifications. Compared to the 2-year import shock, the 5-year import shock is probably better able to reflect the underlying shift in China's import supply capacity.

Table A.1—: China Shock, NTR Voting Records and Electoral Outcomes

$ShareVote_{i,t}$	(1)	(2)	(3)	(4)
$S_{i,t,t-2} \times \text{VoteProCHN}_{i,t,t-2}$	-0.257*	-0.153		
	(0.146)	(0.132)		
$S_{i,t,t-2}$	0.040	-0.013		
	(0.162)	(0.214)		
$\text{VoteProCHN}_{i,t,t-2}$	-0.066**	-0.048**		
	(0.027)	(0.022)		
$S_{i,t,t-5} \times \text{VoteProCHN}_{i,t,t-5}$			-0.150**	-0.103**
			(0.065)	(0.051)
$S_{i,t,t-5}$			0.028	-0.142
			(0.064)	(0.108)
$\text{VoteProCHN}_{i,t,t-5}$			-0.070***	-0.042**
			(0.021)	(0.019)
Additional Controls	Y	Y	Y	Y
State-Year FEs	Y	Y	Y	Y
Individual FEs	N	Y	N	Y
Observations	2,253	2,253	2,253	2,253
R^2	0.525	0.829	0.526	0.829

Notes: Additional controls include: tenure, DW-NOMINATE score, and party affiliation of the politician, and manufacturing employment share in the congressional district and its interaction with NTR voting records. Robust standard errors are two-way clustered at the state and the individual politician level. *** p<0.01, ** p<0.05, * p<0.1

A2. Data generating processes consistent with the rational expectation assumption

In this subsection, we discuss two potential data generating processes that can justify the the rational expectation assumption:

1) In the first scenario, we impose a relatively strong assumption that the agents know the employment shares of different sectors $w_{ik,t}$ but need to predict many sector-level shocks, the exposure to the future shock is then given by:

$$S_{i,t+1} = \sum_k w_{ik,t}\mathbb{E}(S_{k,t+1}|\mathcal{I}_{i,t+1}) + \sum_k w_{ik,t}\varepsilon_{k,t+1}.$$

The expectational errors are specific to k and invariant across i endowed with the same $\mathcal{I}_{i,t}$. Relying on the cross-sectional variation, the rational expectation assumption is then:

$$\sum_k w_{ik,t}\mathbb{E}(\varepsilon_{k,t+1}|\mathcal{I}_{i,t}) = 0. \tag{A.1}$$

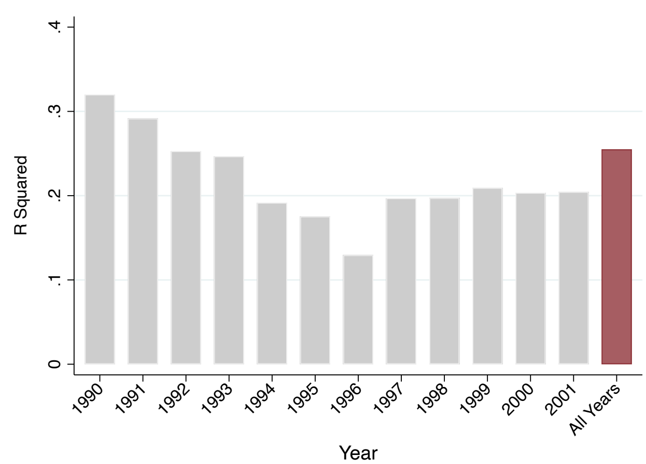
Although $\mathbb{E}(\varepsilon_{k,t+1}|\mathcal{I}_{i,t}) \neq 0$ in general (i.e., agents may over- or under-predict the future shock in a systematic way for a given sector),[42] equation (A.1) only requires that the weighted average of the systematic errors to be zero. What kind of data generating process can support condition (A.1)? Consider the case where politicians endowed with the same information set $\mathcal{I}_{i,t}$ run a weighted regression that autocorrelates the industry-level China shocks across subsequent periods, with weights being the employment shares in the districts, $w_{ik,t}$. (Note that in the context under discussion, the information set $w_{ik,t}$ is in the information set $\mathcal{I}_{i,t}$). The forecast errors then satisfy equation (A.1). This data generating process is plausible, as the representatives have incentives to reduce the forecast errors for the industries with a more important local presence in their district.

Within this context, a potential concern of the validity of rational expectation assumption is that sector-specific shocks are highly correlated. In the extreme case, if the correlation is perfect, agents have only one aggregate shock to predict and (A.1) boils down to $\mathbb{E}(\varepsilon_{t+1}|\mathcal{I}_{i,t}) = 0$. Imposing rational expectation assumption is equivalent to assuming agents have perfect foresight of the future shock, whichs defaults the purpose of our analysis. It turns out that such a correlation can be assessed empirically. We verify that in the data sector-level shocks cannot be subsumed by an aggregate shock (or by a limited number of shocks at the broader sector level). Specifically, we first estimate the variation in the shocks at the broader sector level that

[42]As $\varepsilon_{k,t+1}$ is invariant across i endowed with the same $\mathcal{I}_{i,t}$, if $\mathbb{E}(\varepsilon_{k,t+1}|\mathcal{I}_{i,t}) = 0$, it implies that hat $\varepsilon_{k,t+1} = 0$ for all individuals endowed with $\mathcal{I}_{i,t}$, i.e., agents have perfect foresight on $S_{k,t+1}$.

can be explained by aggregate shocks. Running the regression of sector-level China shocks (i.e., $\frac{M^{oth}_{K,t+5}-M^{oth}_{K,t}}{Y_{K,t}+M_{K,t}-X_{K,t}}$, where K denotes a 2-digit SIC sector) over the period 1990 to 2001 on year fixed effects yields a R-squared of 0.24. We then run regressions of 4-digit SIC-level China shocks (i.e., $\frac{M^{oth}_{k,t+5}-M^{oth}_{k,t}}{Y_{k,t}+M_{k,t}-X_{k,t}}$) on 2-digit SIC fixed effects for each year. If the industry-level shocks are similar within the broader sector, we should expect a large R-squared. The R-squared statistics are reported by the gray bars in Figure A.1, which range from 0.13 to 0.32. Interestingly, the R-squareds are larger in the early 1990s, which echoes our finding in Section 5.2 that the China shocks in the earlier period were more predictable. Finally, as shown by the red bar, the R-squared is 0.25 for the regression that pools all sample periods together and controls for 2-digit SIC $\times$ year fixed effects. In sum, there is substantial residual variation in shocks across broader sectors, and even across disaggregated industries within these broader sectors.

Figure A.1. : The Variation of 4-digit SIC-Level China Shocks Explained by 2-digit SIC Fixed Effects



2) The discussion so far has relied on the assumption that agents perfectly know the weights $w_{ik,t}$, which is a stringent requirement. While it is plausible that politicians know the manufacturing employment share in their districts, having the exact knowledge of current employment share for each sector requires the most up-to-date data which may not be available when the most NTR voting decisions are made. There are again two potential data generating process for this case.

The first data generating process has politicians endowed with $\mathcal{I}_{i,t}$ forming expectation for $w_{ik,t}$ and $S_{k,t+1}$ separately. That is

$$w_{ik,t} = \mathbb{E}(w_{ik,t}|\mathcal{I}_{i,t}) + \varepsilon^{w}_{ik,t} \quad \text{and} \quad S_{k,t+1} = \mathbb{E}(S_{k,t+1}|\mathcal{I}_{i,t}) + \varepsilon^{s}_{k,t+1}.$$

In this case, the expected future exposure that affects the voting decision through the lens of our model is $\sum_k \mathbb{E}(w_{ik,t}|\mathcal{I}_{i,t})\mathbb{E}(S_{k,t+1}|\mathcal{I}_{i,t})$. As with our baseline case, researchers do not observe agents' information sets and hence expectations, and need to replace the term by $S_{i,t+1} = \sum_k w_{ik,t}S_{k,t+1}$. It is straightforward to show that if (i) $\sum_k \mathbb{E}(w_{ik,t}\varepsilon^{s}_{k,t+1}|\mathcal{I}_{i,t}) = 0$, and (ii) $\mathbb{E}(\varepsilon^{w}_{ik,t}|\mathcal{I}_{i,t}) = 0\ \forall k$, all the moment inequality conditions still hold. Condition (i) is analogous to equation (A.1), and condition (ii) imposes the rational expectation assumption for $w_{ik,t}$. These averages are taken across individuals. To see this:

$$\begin{aligned} S_{i,t+1} &= \sum_k w_{ik,t}S_{k,t+1} \\ &= \sum_k \mathbb{E}(w_{ik,t}|\mathcal{I}_{i,t})\mathbb{E}(S_{k,t+1}|\mathcal{I}_{i,t}) + \underbrace{\sum_k w_{ik,t}\varepsilon^{s}_{k,t+1} + \sum_k \varepsilon^{\omega}_{ik,t}\mathbb{E}(S_{k,t+1}|\mathcal{I}_{i,t})}_{\nu_{i,t+1}}. \end{aligned}$$

When conditions (i) and (ii) hold, we have $\mathbb{E}(\nu_{i,t+1}|\mathcal{I}_{i,t}) = 0$ and can invoke Jensen's inequality to derive all the moment inequality conditions.

The second data generating process has the politicians forming expectations on $S_{i,t+1} = \sum_k w_{ik,t}S_{k,t+1}$ as a whole. Then, rational expectation only requires that $\mathbb{E}(\varepsilon_{i,t+1}|\mathcal{I}_{i,t}) = 0$ where $\varepsilon_{i,t+1} = S_{i,t+1} - \mathbb{E}(S_{k,t+1}|\mathcal{I}_{i,t})$.

A3. *Role of expectations about electoral sensitivity (γ)*

Our baseline model assumes that politicians have full knowledge of the electoral consequences of the China shock that are dependent on the NTR voting decisions. In this appendix, we relax this assumption. Denote $\gamma_t = \gamma^1_t - \gamma^2_t$ as the electoral sensitivity to the China shock. Suppose politicians form expectation on both γ_t and $S_{i,t+1}$. Then the voting decision is determined by

$$Y_{i,t} = \mathbb{1}\{a_t\theta_i + b_t + \tilde{\delta}\mathbb{E}[\gamma_t S_{i,t+1}|\mathcal{I}_{i,t}] - \xi_{i,t} > 0\}.$$

Here we make two assumptions: (i) γ_t and $S_{i,t+1}$ are uncorrelated conditional on $\mathcal{I}_{it}$, and (ii) the distribution of γ_t conditional on $\mathcal{I}_{it}$ is equal to its distribution conditional on an information set common to all politicians at time t, $\mathcal{I}_t$. Under these assumptions we can rewrite the decision rule as:

$$Y_{i,t} = \mathbb{1}\{a_t\theta_i + b_t + \tilde{\delta}\mathbb{E}\left[\gamma_t|\mathcal{I}_t\right]\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}] - \xi_{i,t} > 0\}.$$

It is straightforward to redefine the coefficient $\delta_t = \tilde{\delta}\mathbb{E}\left[\gamma_t|\mathcal{I}_t\right]$ and reinterpret all the results we report under this slightly different definition. We believe assumptions (i) and (ii) are not implausible because the sensitivity of voting to economic shocks is a common parameter that does not depend on the individual politician's characteristics. We see γ_t as a parameter that politicians estimate from voter surveys on the importance of certain issues in an election.[43] The independence of the China shock from this parameter γ_t is also reasonable once we consider that this is a common parameter that does not depend on specific constituencies.

[43]See Jones and Baumgartner (2005).

Monte Carlo simulation: MLE bias

In this Appendix, we adopt Monte Carlo simulations to illustrate the bias that emerges when estimating ω_t by maximum likelihood in equation (7). We adopt a time horizon of 5 years to define a future import shock, so that $S_{i,t+1}$ corresponds to the increase in import penetration in the next 5 years. The true information set throughout the exercise is what we have defined as baseline information set $\mathcal{I}^b_{i,t} = \{S_{it,}, ShareMfg_{i,t}, \theta_i\}$. More specifically, $S_{i,t+1}$ is simulated according to the following linear model:

$$S_{i,t+1} = \underbrace{\beta_0 + \beta_1 S_{i,t} + \beta_2 ShareMfg_{i,t} + \beta_3 \theta_i}_{\mathbb{E}[S_{i,t+1}|\mathcal{I}^b_{i,t}]} + \varepsilon_{i,t+1},$$

where $\varepsilon_{i,t+1}$ is the expectational error. To simulate $S_{i,t+1}$, we set $\beta_0 = 0$, $\beta_1 = 0.721$, $\beta_2 = 0.184$, $\beta_3 = 0$. Expectational errors $\varepsilon_{i,t+1}$ are drawn from the normal distribution with mean 0 and standard deviation of $\sigma_\varepsilon = 0.525$.[44] We take $S_{i,t}$, $ShareMfg_{i,t}$ and θ_i from the real data over 1998-2001. Then, the voting decision is simulated according to $Y_{i,t} = \mathbb{1}\{a\theta_i + b + \delta\mathbb{E}[S_{i,t+1}|\mathcal{I}^b_{i,t}] - \xi_{i,t} > 0\}$ with $\xi_{i,t} \sim N(0,1)$. We set $a = 0.5$ and $b = 0.3$ and present two sets of results based on $\delta = -1.3$ and $\delta = 0$.

We will evaluate the bias that arises when an econometrician mistakenly assumes that politicians have minimal information $\mathcal{I}^m_{i,t}$ or perfect foresight $\mathcal{I}^p_{i,t}$. For the case of minimal information set, we estimate ω_t by maximizing the following log-likelihood:

$$\begin{aligned} \ln\mathcal{L}\left(\omega_t | \{Y_{i,t}, \theta_i, \mathcal{I}_{i,t}\}_{i=1}^N\right) = & \sum_{i=1}^N Y_{i,t} \ln\left[\Phi\left(a_t\theta_i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^m_{i,t}\right]\right)\right] \\ & + (1 - Y_{i,t}) \ln\left[1 - \Phi\left(a_t\theta_i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^m_{i,t}\right]\right)\right]. \end{aligned} \tag{B.1}$$

where the expectation is estimated as:

$$S_{i,t+1} = \underbrace{\mu_0 + \mu_1 ShareMfg_{i,t} + \mu_2 \theta_i}_{\mathbb{E}[S_{i,t+1}|\mathcal{I}^m_{i,t}]} + u_{i,t+1}$$

For the case of perfect foresight the log-likelihood is maximized after replacing the expected future shocks with $S_{i,t+1}$.

Table B.1 reports the mean and the standard deviation of MLE estimates for 500 simulated datasets. First, the simulations clearly indicate that when the model is correctly specified, the average parameter estimates are very close to the

[44] The parameters β's are estimated based on the actual data over 1998-2001. σ_ε is set based on the empirical distribution of the regression residuals.

true parameters. Second, when we assume the politician has perfect foresight, i.e., more information than she actually does, there is a clear attenuation bias. The average $\hat{\delta}$ is -0.813 (the true δ is -1.3). We discussed the intuition for this attenuation bias as related to classical measurement error in the main text. A more nuanced case is the one in which the politician is assumed to have a minimal information set, i.e., the econometrician assumes that the politician knows strictly less than what she actually knows. In our simulations we find again an attenuation bias, as the average $\hat{\delta}$ is -1.060. In our example the true voting decision is taken based on the baseline information set, as follows:

$$\begin{aligned} Y_{i,t} &= 1\{a_t\theta_i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^b_{i,t}\right] - \xi_{i,t} \geq 0\} \\ &= 1\{a_t\theta_i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^m_{i,t}\right] + \delta_t \left\{\mathbb{E}\left[S_{i,t+1}|\mathcal{I}^b_{i,t}\right] - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^m_{i,t}\right]\right\} - \xi_{i,t} \geq 0\}, \end{aligned}$$

where the error term becomes $\rho_{i,t} = \delta_t \left\{\mathbb{E}\left[S_{i,t+1}|\mathcal{I}^b_{i,t}\right] - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^m_{i,t}\right]\right\} - \xi_{i,t}$, with a standard deviation of $\sigma_\rho > \sigma_\xi = 1$. It is straightforward to show that $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}^b_{i,t}\right] - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^m_{i,t}\right]$ and $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}^m_{i,t}\right]$ are uncorrelated by construction. That is, the measurement error is uncorrelated with the assumed proxy of agents' expectation (which is different from the case of perfect foresight). However, in this case, the MLE estimator rescales the true parameter by $1/\sigma_\rho < 1$, which still results in an attenuation bias. (Yatchew and Griliches, 1985)

Finally, we consider the special case in which $\delta = 0$. Clearly, if the China shock truly does not matter for voting decisions, MLE biases are negligible.

Table B.1—: Maximum Likelihood - Monte Carlo Simulation

	Assumed Information Set	Avg $\hat{a}$ (std.)	Avg $\hat{b}$ (std.)	Avg $\hat{\delta}$ (std.)
$a = 0.5, b = 0.3, \delta = -1.3$	(1) Minimal Information	0.449 (0.066)	0.303 (0.027)	-1.060 (0.047)
	(2) Baseline Information (correct)	0.498 (0.079)	0.319 (0.034)	-1.306 (0.058)
	(3) Perfect Foresight	0.421 (0.090)	0.304 (0.040)	-0.813 (0.190)
$a = 0.5, b = 0.3, \delta = 0$	(4) Minimal Information	0.499 (0.073)	0.300 (0.029)	-0.001 (0.046)
	(5) Baseline Information (correct)	0.499 (0.072)	0.300 (0.029)	-0.000 (0.036)
	(6) Perfect Foresight	0.500 (0.072)	0.300 (0.029)	-0.002 (0.041)

Estimation strategy, specification tests and counterfactual simulations: details

C1. Estimation implementation

Conditional moments (10), (11), (14) and (16) cannot be directly employed for empirical applications because conditioning on each possible value of $Z_{i,t}$ is computationally unfeasible. The standard solution in the moment inequality literature, which we adopt, is to transform conditional into unconditional moment inequalities, which can be directly employed in estimation.[45] This is not innocuous in that information is lost in transitioning from conditional inequalities to a relatively smaller set of unconditional inequalities. As a result, the parameters that satisfy conditional moment inequalities may be a small subset of those that satisfy the unconditional moments. Whether these larger confidence sets remain sufficiently informative is an empirical question.[46]

We collect the four moment inequalities (10), (11), (14) and (16) and we adopt the unconditional moment inequalities:

$$\mathbb{E}\left[\left\{\begin{array}{c} m_l^{ob} \\ m_u^{ob} \\ m_l^{rp} \\ m_u^{rp} \end{array}\right\} \times g\left(Z_{i,t}\right)\right] \geq 0, \tag{C.1}$$

where the instrument function $g\left(Z_{i,t}\right)$ is specified, e.g., for the minimal information case of $Z_{i,t} = \{\theta_i, ShareMfg_{i,t}\}$, as follows:

$$g\left(Z_{i,t}\right) = \left\{\begin{array}{c} \mathbf{1}\left\{\theta_i > med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{ShareMfg_{i,t} > med\left(ShareMfg_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i > med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{ShareMfg_{i,t} \leq med\left(ShareMfg_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i \leq med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{ShareMfg_{i,t} > med\left(ShareMfg_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i \leq med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{ShareMfg_{i,t} \leq med\left(ShareMfg_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i > med\left(\theta_i\right)\right\} \times \left|\theta_i - med\left(\theta_i\right)\right| \\ \mathbf{1}\left\{\theta_i \leq med\left(\theta_i\right)\right\} \times \left|\theta_i - med\left(\theta_i\right)\right| \\ \mathbf{1}\left\{ShareMfg_{i,t} > med\left(ShareMfg_{i,t}\right)\right\} \times \left|ShareMfg_{i,t} - med\left(ShareMfg_{i,t}\right)\right| \\ \mathbf{1}\left\{ShareMfg_{i,t} \leq med\left(ShareMfg_{i,t}\right)\right\} \times \left|ShareMfg_{i,t} - med\left(ShareMfg_{i,t}\right)\right| \end{array}\right. .$$

Instead of conditioning on all the possible values of θ_i and $ShareMfg_{i,t}$, this approach calculates for example the moment inequalities in different quardrants of the space consitituted by θ_i and $ShareMfg_{i,t}$. For the minimal information case, we have $8 \times 4 = 32$ moment inequalities, which we use to construct confidence sets, as explained in section IV.[47] The choice of the instrument functions does

[45]Starting from a moment inequality of the form $\mathbb{E}[m|Z] \geq 0$, let us consider an instrument function $g\left(Z\right) > 0$. Multiplying the two and taking expectation yields $\mathbb{E}\left[g\left(Z\right)\mathbb{E}[m|Z]\right] \geq 0$ which implies $\mathbb{E}\left[g\left(Z\right)m\right] \geq 0$ whenever $g\left(Z\right)$ is Z-measurable.

[46]For a complete discussion see Andrews and Shi (2013).

[47]For the case of baseline information $Z_{i,t} = \{\theta_i, ShareMfg_{i,t}, S_{i,t}\}$, the instrument function is given

not appear to drive our results and we probed it in several robustness checks. For instance, results remain qualitatively similar when we limit the analysis to the subset of $g\left(Z_{i,t}\right)$ containing only the pairwise interactions of relevant dummy variables.

C2. Confidence sets for parameters

The modified method of moment test statistics is:

$$Q(\omega_p) = \sum_{k=1}^{K} \left(\min\{ \frac{\bar{m}_k(\omega_p)}{\hat{\sigma}_k(\omega_p)}, 0\} \right)^2 \tag{C.2}$$

where ω_p is a point in the space Ω_g, and K is the number of moments. $\bar{m}_k(\omega_p) = \frac{1}{n}\sum_i \sum_t m_k(S_{i,t+1}, \theta_i, \omega_p)$ is the mean value of the moment k evaluated at ω_p, and $\hat{\sigma}_k(\omega_p)$ is the corresponding standard error. When $Q(\omega_p) = 0$, it indicates that all the moment listed in (C.1) are satisfied at ω_p, and hence ω_p could be included in the identified set. If $Q(\omega_p) > 0$, it indicates that some sample moment inequalities are violated when evaluated at ω_p. This may result from two independent cases: (i) some population moment inequalities are indeed violated at ω_p; and (ii) some sample moment inequalities are violated because of sampling variation (Ho and Pakes, 2014).

To account for the sampling variation, we adopt the Generalized Moment Selection (GMS) test in Andrews and Soares (2010) which simulates the asymptotic distribution of $Q(\omega_p)$ under the null hypothesis $H_o : \omega^* = \omega_p$. Here, ω^* denotes the true parameter vector. More specifically, the simulation is based on R draws

by:

$$g\left(Z_{i,t}\right) = \begin{cases} \mathbf{1}\left\{\theta_i > med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{ShareMfg_{i,t} > med\left(ShareMfg_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i > med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{ShareMfg_{i,t} \leq med\left(ShareMfg_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i \leq med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{ShareMfg_{i,t} > med\left(ShareMfg_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i \leq med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{ShareMfg_{i,t} \leq med\left(ShareMfg_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i > med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{S_{i,t} > med\left(S_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i > med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{S_{i,t} \leq med\left(S_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i \leq med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{S_{i,t} > med\left(S_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i \leq med\left(\theta_i\right)\right\} \times \mathbf{1}\left\{S_{i,t} \leq med\left(S_{i,t}\right)\right\} \\ \mathbf{1}\left\{ShareMfg_{i,t} > med\left(ShareMfg_{i,t}\right)\right\} \times \mathbf{1}\left\{S_{i,t} > med\left(S_{i,t}\right)\right\} \\ \mathbf{1}\left\{ShareMfg_{i,t} > med\left(ShareMfg_{i,t}\right)\right\} \times \mathbf{1}\left\{S_{i,t} \leq med\left(S_{i,t}\right)\right\} \\ \mathbf{1}\left\{ShareMfg_{i,t} \leq med\left(ShareMfg_{i,t}\right)\right\} \times \mathbf{1}\left\{S_{i,t} > med\left(S_{i,t}\right)\right\} \\ \mathbf{1}\left\{ShareMfg_{i,t} \leq med\left(ShareMfg_{i,t}\right)\right\} \times \mathbf{1}\left\{S_{i,t} \leq med\left(S_{i,t}\right)\right\} \\ \mathbf{1}\left\{\theta_i > med\left(\theta_i\right)\right\} \times \left|\theta_i - med\left(\theta_i\right)\right| \\ \mathbf{1}\left\{\theta_i \leq med\left(\theta_i\right)\right\} \times \left|\theta_i - med\left(\theta_i\right)\right| \\ \mathbf{1}\left\{ShareMfg_{i,t} > med\left(ShareMfg_{i,t}\right)\right\} \times \left|ShareMfg_{i,t} - med\left(ShareMfg_{i,t}\right)\right| \\ \mathbf{1}\left\{ShareMfg_{i,t} \leq med\left(ShareMfg_{i,t}\right)\right\} \times \left|ShareMfg_{i,t} - med\left(ShareMfg_{i,t}\right)\right| \\ \mathbf{1}\left\{S_{i,t} > med\left(S_{i,t}\right)\right\} \times \left|S_{i,t} - med\left(S_{i,t}\right)\right| \\ \mathbf{1}\left\{S_{i,t} > med\left(S_{i,t}\right)\right\} \times \left|S_{i,t} - med\left(S_{i,t}\right)\right| \end{cases}.$$

Hence, there are in total $18 \times 4 = 64$ moment inequalities. For the case of perfect foresight $Z_{i,t} = \{\theta_i, ShareMfg_{i,t}, S_{i,t}, \varepsilon_{i,t+1}\}$ where $\varepsilon_{i,t+1}$ captures all the information orthogonal to the elements in the baseline information set. The instrument function is augmented accordingly, and there are in total 128 moment inequality conditions.

from the multivariate normal distribution $N(0_K, I_K)$, where I_K is the identity matrix of dimension K. Each draw ζ_r yields the following statistics:

$$Q_r^{AA}(\omega_p) = \sum_{k=1}^{K} \left(\left(\min\{[\hat{\Lambda}^{\frac{1}{2}}(\omega_p)\zeta_r]_k, 0\} \right)^2 \times \mathbb{1}\{\sqrt{n}\frac{\bar{m}_k(\omega_p)}{\hat{\sigma}_k(\omega_p)} \leq \sqrt{\ln n}\} \right) \tag{C.3}$$

where $\hat{\Lambda}(\cdot)$ is the estimate of the variance-covariance matrix of $\frac{m_k(S_{i,t+1},\theta_i,\omega_p)}{\hat{\sigma}_k(\omega_p)}$. $[\hat{\Lambda}^{\frac{1}{2}}(\omega_p)\zeta_r]_k$ is the kth element of the vector $\hat{\Lambda}^{\frac{1}{2}}(\omega_p)\zeta_r$, and it is the simulated counterpart of $\frac{\bar{m}_k(\omega_p)}{\hat{\sigma}_k(\omega_p)}$ in equation (C.2).[48] Pooling all $Q_r^{AA}(\omega_p)$ together renders the simulated distribution denoted by $Q^{AA}(\omega_p)$.

Let $\hat{c}^{AA}(\omega_p, 1-\alpha)$ denote the $(1-\alpha)$-quantile of $Q^{AA}(\omega_p)$. ω_p is included in the $(1-\alpha)\%$ confidence set (CS) if $nQ(\omega_p) \leq \hat{c}^{AA}(\omega_p, 1-\alpha)$. By repeating the procedure for each point ω_p in the grid space Ω_p, we derive the $(1-\alpha)\%$ CS denoted by $\hat{\Omega}^{1-\alpha}$.[49] In the main text, we set $\alpha = 0.05$, and report the 95% confidence sets.

C3. BP, RC, and RS test statistics and corresponding p-values

This appendix details the statistical tests of whether covariates in $Z_{i,t}$ are contained in the information set possessed by politicians when they vote on the NTR with China. Intuitively, when the original voting model is correctly specified, but the information set is misspecified by researchers, i.e $Z_{i,t} \notin \mathcal{I}_{i,t}$, some moment inequalities will be violated and the confidence set is likely to be empty. In the following, we discuss the BP, RS and RS tests based on Bugni, Canay and Shi (2015).

P-value of the Test BP

The test statistics for the Test BP is defined as in equation (C.2). For a given value of λ, we infer whether $\forall\, \omega_p \in \Omega_g$, $Q(\omega_p) > \hat{c}^{AA}(\omega_p, 1-\lambda)$. If so, we lower λ by a small amount, and repeat the procedure until reaching the value λ^{BP} such that $\exists\, \omega_p \in \Omega_g$ such that $Q(\omega_p) \leq \hat{c}^{AA}(\omega_p, 1-\lambda^{BP})$. λ^{BP} is then the p-value for the test BP.[50] As pointed out by Bugni, Canay and Shi (2015), this test is too conservative. Therefore, we turn to the test RC and the test RS proposed by the authors as follows.

[48]The component $\mathbb{1}\{\sqrt{n}\frac{\bar{m}_k(\omega_p)}{\hat{\sigma}_k(\omega_p)} \leq \sqrt{\ln n}\}$ is the generalized moment selection function which selects the moment that are almost binding.

[49]We conduct the grid search within a predefined grid space Ω_g. If some of the points in $\hat{\Omega}^{1-\alpha}$ are at the boundary of Ω_g, we expand the limits of the grid space and repeat the procedure described above. For our baseline model, we fill the 3-dimensional space with 64,000 equidistant grids. For the augmented model in section IV.D and Appendix E.E3, we fill the 4-dimensional space with 160,000 equidistant grids.

[50]In practice, we start the algorithm with $\lambda = 0.99$ and reduce it by 0.005 at a time. We stop when λ reaches 0.01. Hence, p-value=0.01 (respectively, 0.99) in our tables indicates that the p-value is less than or equal to 0.01 (respectively, greater than or equal to 0.99).

P-value of the Test RC

To conduct the test RC, we first calculate the minimum of test statistics (C.2) across all $\omega_p \in \Omega_g$. Denote the minimum by $T = \min_{\omega \in \Omega_g} Q(\omega_p)$, and $\Omega_g^* = \arg\min_{\Omega_p}(Q(\omega_p))$.[51] The p-value of the test RC is constructed as follows. For a given value of λ, we compute $\hat{c}^{RC}(1-\lambda) = \min_{\omega_p \in \Omega_g^*} \hat{c}^{AA}(\omega_p, 1-\lambda)$. If $T > \hat{c}^{RC}(1-\lambda)$, we lower λ by a small amount, and repeat the procedure until reaching the value λ^{RC} such that $T \leq \hat{c}^{RC}(1-\lambda^{RC})$. Then, λ^{RC} is the corresponding p-value.

P-value of the Test RS

Similar to the test RC, to conduct the test RS, we first compute T and derive Ω_g^* as defined above. Then, we use the R draws from the multivariate normal distribution $N(0_K, I_K)$. For each draw ζ_r, we compute $Q_r^{AA}(\omega_p)$ for each $\omega_p \in \Omega_g^*$ according to equation (C.3), and derive the corresponding minimum $T_r = \min_{\omega_p \in \Omega_g^*} Q_r^{AA}(\omega_p)$. The p-value of the test RS is constructed as follows. For a given value of λ, we find the $(1-\lambda)$-quantile of T_r, denoted by $\hat{c}^{RS}(1-\lambda)$. If $T > \hat{c}^{RS}(1-\lambda)$, we lower λ by a small amount, and repeat the procedure until reaching the value λ^{RS} such that $T \leq \hat{c}^{RS}(1-\lambda^{RS})$. Then, λ^{RS} is the corresponding p-value.

Note that the test RC and the test RS are different in the following way. For the test RC, to derive the critical value $\hat{c}^{RC}(1-\lambda)$, we first compute the $(1-\lambda)$-quantile for each asymptotic distributions Q^{AA} evaluated at each $\omega_p \in \Omega_g^*$, and then take the minimum across these quantiles. For the RS test, to derive the critical value $\hat{c}^{RS}(1-\lambda)$, we first compute the minimum of $Q_r^{AA}(\omega_p)$ across $\omega_p \in \Omega_g^*$, and then derive the $(1-\lambda)$-quantile of these minimums. When Ω_g^* contains only one point, $\hat{c}^{RC}(1-\lambda) = \hat{c}^{RS}(1-\lambda)$, and the p-values from the two tests are the same. When Ω_g^* contains multiple points, $\hat{c}^{RC}(1-\lambda) \geq \hat{c}^{RS}(1-\lambda)$, and the p-value of the test RS will be no larger than that of the test RC.

C4. Counterfactual simulations: details

Conditional on a particular value of the parameter vector $\omega_t = \{a_t, b_t, \delta_t\}$, and information set $\mathcal{I}_{i,t}$, the decision of voting in favor of China is given by:

$$Y_{i,t}(\omega_t, \mathcal{I}_{i,t}, \xi_{i,t}) = \{a_t\theta_i + b_t + \delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}] - \xi_{i,t} \geq 0\}.$$

[51] As discussed in Ho and Pakes (2014) (pg. 3868), Ω_g^* could be a set of values all of which make (C.2) zero (i.e., all the moments are satisfied), or it could be a point, indicating that no value of ω satisfies all the moment conditions. The latter case could be a result of sampling error, which is accounted for by the GMS approach proposed by Andrews and Soares (2010). In our case, Ω_g^* contains only one point as the case in Ho and Pakes (2014).

Integrating $Y_{i,t}(\omega_t, \mathcal{I}_{i,t}, \xi_{i,t})$ over $\xi_{i,t}$ generates the probability that politician i casts a pro-China vote in period t. In particular,

$$\int_{\xi_{i,t}} Y_{i,t}(\omega_t, \mathcal{I}_{i,t}, \xi_{i,t})\phi(\xi_{i,t})d\xi_{i,t} = \int_{\xi_{i,t}} \mathbf{1}\left\{a_t\theta_i + b_t + \delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}] - \xi_{i,t} \geq 0\right\}\phi(\xi_{i,t})d\xi_{i,t}.$$

Denote by $\mathcal{N}_t$ the set of politicians in period t, the share of votes in favor of China is then given by:

$$\Pi^+(\omega_t, \mathcal{I}_{i,t}, \mathcal{N}_t) = \frac{1}{N_t}\sum_{i\in\mathcal{N}_t}\int_{\xi_{i,t}} Y_{i,t}(\omega_t, \mathcal{I}_{i,t}, \xi_{i,t})\phi(\xi_{i,t})d\xi_{i,t}.$$

The corresponding 95 percent confidence set of the number of politicians vote in favor of China is:

$$\left[\min_{\omega_t\in\Omega_t^{95\%}}\left\{\Pi^+(\omega_t, \mathcal{I}_{i,t}, \mathcal{N}_t)\right\}, \max_{\omega_t\in\Omega_t^{95\%}}\left\{\Pi^+(\omega_t, \mathcal{I}_{i,t}, \mathcal{N}_t)\right\}\right],$$

where $\Omega_t^{95\%}$ is the 95 percent confidence set for the underlying parameters.

Given the assumption of normal distribution and with a large sample, the change of vote share in favor of NTR with China when information is improved from $\mathcal{I}_{i,t}^b$ to $\mathcal{I}_{i,t}^p$ can be written as:

$$\begin{aligned}
& \Pi^+(\omega_t, \mathcal{I}_{i,t}^p, \mathcal{N}_t) - \Pi^+(\omega_t, \mathcal{I}_{i,t}^b, \mathcal{N}_t) \\
= & \frac{1}{N_t}\sum_{i\in\mathcal{N}_t}[\Phi(a_t\theta_i + b_t + \delta_t S_{i,t+1}) - \Phi(a_t\theta_i + b_t + \delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b])] \\
= & \int_{\varepsilon_{i,t+1}}[\Phi(a_t\theta_i + b_t + \delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b] + \delta_t\varepsilon_{i,t+1}) - \Phi(a_t\theta_i + b_t + \delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b])]dG(\varepsilon_{i,t+1}).
\end{aligned}$$

The corresponding 95% confidence set is given by:

$$\left[\min_{\omega_t\in\Omega_t^{95\%}}\left\{\Pi^+(\omega_t, \mathcal{I}_{i,t}^p, \mathcal{N}_t) - \Pi^+(\omega_t, \mathcal{I}_{i,t}^b, \mathcal{N}_t)\right\}, \max_{\omega_t\in\Omega_t^{95\%}}\left\{\Pi^+(\omega_t, \mathcal{I}_{i,t}^p, \mathcal{N}_t) - \Pi^+(\omega_t, \mathcal{I}_{i,t}^b, \mathcal{N}_t)\right\}\right].$$

We also simulate the number of politicians who vote for China in the baseline, but switch vote in the counterfactual according to:

$$\begin{aligned}
& N^{+-}(\omega_t, \mathcal{I}_{i,t}^b \to \mathcal{I}_{i,t}^p, \mathcal{N}_t) \\
= & \sum_{i\in\mathcal{N}_t}\int_{\xi_{i,t}} \mathbf{1}\left\{a_t\theta_i + b_t + \delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b] - \xi_{i,t} \geq 0\right\}\mathbf{1}\left\{a_t\theta_i + b_t + \delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^p] - \xi_{i,t} < 0\right\}\phi(\xi_{i,t})d\xi_{i,t}.
\end{aligned}$$

Similarly, the number of politicians who vote in favor of China in both the baseline and the counterfactual is:

$$N^{++}(\omega_t, \mathcal{I}^b_{i,t} \to \mathcal{I}^p_{i,t}, \mathcal{N}_t)$$
$$= \sum_{i \in \mathcal{N}_t} \int_{\xi_{i,t}} \mathbf{1}\left\{a_t\theta_i + b_t + \delta_t \mathbb{E}[S_{i,t+1}|\mathcal{I}^b_{i,t}] - \xi_{i,t} \geq 0\right\} \mathbf{1}\left\{a_t\theta_i + b_t + \delta_t \mathbb{E}[S_{i,t+1}|\mathcal{I}^p_{i,t}] - \xi_{i,t} \geq 0\right\} \phi(\xi_{i,t}) d\xi_{i,t}.$$

The number of politicians who vote against China in the baseline and switch in the counterfactual is:

$$N^{-+}(\omega_t, \mathcal{I}^b_{i,t} \to \mathcal{I}^p_{i,t}, \mathcal{N}_t)$$
$$= \sum_{i \in \mathcal{N}_t} \int_{\xi_{i,t}} \mathbf{1}\left\{a_t\theta_i + b_t + \delta_t \mathbb{E}[S_{i,t+1}|\mathcal{I}^b_{i,t}] - \xi_{i,t} < 0\right\} \mathbf{1}\left\{a_t\theta_i + b_t + \delta_t \mathbb{E}[S_{i,t+1}|\mathcal{I}^p_{i,t}] - \xi_{i,t} \geq 0\right\} \phi(\xi_{i,t}) d\xi_{i,t}.$$

The number of politicians who vote against China in both the baseline and the counterfactual is:

$$N^{--}(\omega_t, \mathcal{I}^b_{i,t} \to \mathcal{I}^p_{i,t}, \mathcal{N}_t)$$
$$= \sum_{i \in \mathcal{N}_t} \int_{\xi_{i,t}} \mathbf{1}\left\{a_t\theta_i + b_t + \delta_t \mathbb{E}[S_{i,t+1}|\mathcal{I}^b_{i,t}] - \xi_{i,t} < 0\right\} \mathbf{1}\left\{a_t\theta_i + b_t + \delta_t \mathbb{E}[S_{i,t+1}|\mathcal{I}^p_{i,t}] - \xi_{i,t} < 0\right\} \phi(\xi_{i,t}) d\xi_{i,t}.$$

The share of politicians for each of the cases, and the corresponding confidence sets are defined accordingly.

Data appendix

D1. Trade, employment and output data

This subsection describes the data sources that we employ to construct the commuting-zone-level measures (18) and (19). Data on bilateral trade flows over the period 1988-2006 for the 4-digit Standard International Trade Classification (SITC) industries are obtained from UN Comtrade Database. [52] We concord these data to four-digit Standard Industry Classification (SIC) industries.[53] Following Autor, Dorn and Hanson (2013) and Acemoglu et al. (2016), we aggregate together a few four-digit industries to ensure compatibility with the additional data on employment discussed below. Our final data set contains 397 manufacturing industries. We complement the trade data by the output data obtained from the NBER-CES data.[54] All import, export and output amounts are inflated to 2007 US dollars using the Personal Consumption Expenditure (PCE) deflator.

Information on industry employment structure by CZs over 1988-2001 is derived from the County Business Patterns (CBP) data published by the US Census Bureau.[55] The CBP tracks employment, firm size distribution, and payroll by county and industry annually. To protect confidentiality, employment for county-industry cells is sometimes reported as an interval instead of exact count. We use the fixed-point imputation algorithm developed by Autor, Dorn and Hanson (2013) to derive employment for each county-industry cell. The county-level data is then aggregated to commuting zones using the concordances provided by Autor, Dorn and Hanson (2013).[56]

D2. NTR votes for Vietnam: additional details

Roll call votes on NTR with Vietnam. We collect data on voting outcomes of bills related to the renewal of Vietnam's Jackson-Vanik Waiver (i.e., whether to extend Vietnam's NTR status) that existed over the period 1998 to 2002. Due to the congressional redistricting in 2002, we only include the bills over 1998-2001. These bills are HJRES120, HJRES58, HJRES99 and HJRES55. The shares of

[52]See https://wits.worldbank.org.

[53]The crosswalk that cross-matches the four-digit SITC (Rev.2) industries and the four-digit SIC (1987 version) is constructed as follows. (1) We first map the four-digit SITC industries to the corresponding six-digit Harmonized Commodity Description and Coding System (HS) products based on the concordance provided by UN WITS (https://wits.worldbank.org/product_concordance.html). (2) We then apply the crosswalk from Autor, Dorn and Hanson (2013), which assigns 6-digit HS products to 4-digit SIC industries. (3) Lastly, the four-digit SITC codes are cross-matched with the four-digit SIC codes based on their relations with the six-digit HS codes. We complement the data with the US SIC87-level imports and exports data (1972-2005) from Shott's web site (https://sompks4.github.io/sub_data.html).

[54]See https://www.nber.org/research/data/nber-ces-manufacturing-industry-database.

[55]The repository of the data over the sample period is the United States National Archives (https://catalog.archives.gov/id/613576).

[56]Industry classifications in CBP changed periodically — over 1988-1997, employment is classified using the SIC (1987 version) codes, while employment thereafter is expressed according to the North American Industry Classification System (NAICS). Using the crosswalk in Autor, Dorn and Hanson (2013), we concord the post-1997 data to the four-digit SIC industries.

votes in favor of NTR with Vietnam are 61.47%, 69.63%, 78.54% and 78.12% in 1998, 1999, 2000 and 2001, respectively.

Import shock from Vietnam. Analogous to the China shock, we construct the future import supply shocks from Vietnam at the CZ level according to

$$S^V_{j,t+1} = \sum_k \frac{L_{jk,t}}{L_{j,t}} \frac{\Delta M^{oth,V}_{k,t+1}}{Y_{k,t} + M_{k,t} - X_{k,t}},$$

where $\Delta M^{oth,V}_{k,t+1}$ is the change in import of good k from Vietnam by eight other (non-U.S.) high-income countries over 5 years in the future. The past import shock from Vietnam is

$$S^V_{j,t} = \sum_k \frac{L_{ik,t-5}}{L_{i,t-5}} \frac{\Delta M^{oth,V}_{k,t}}{Y_{k,t-5} + M_{k,t-5} - X_{k,t-5}},$$

where $\Delta M^{oth,V}_{k,t}$ denotes the change in import of good k from Vietnam by eight other (non-U.S.) high-income countries over the past 5 years. We aggregate the CZ-level measures to congressional districts using the same procedure in section III.B. The CD-level measures are denoted by $S^V_{i,t+1}$ and $S^V_{i,t}$. The magnitude of the import shock from Vietnam is several orders smaller than that from China. The standard deviation of $S^V_{i,t+1}$ over 1998-2001 is 0.0057. As with China, local manufacturing share and past import shock have a large predicting power for future shock. Regressing (de-trended) $S^V_{i,t+1}$ on (de-trended) $S^V_{i,t}$ and $ShareMfg_{i,t}$ yields a R-squared of 0.528.

Table D.1—: Roll Call Votes

Year	Congress	President	House	Bill number	NTR (PNTR) approved in House	Additional action
Annual Renewal of NTR with China:						
1990	101	G.H.W. Bush	D	HJRES647	No	No action in Senate
1991	102	G.H.W. Bush	D	HJRES263	No	No action in Senate
1992	102	G.H.W. Bush	D	HJRES502	No	Did not pass in Senate
1993	103	Clinton	D	HJRES208	Yes	
1994	103	Clinton	D	HJRES373	Yes	
1995	104	Clinton	R	HJRES96	Yes	
1996	104	Clinton	R	HJRES182	Yes	
1997	105	Clinton	R	HJRES79	Yes	
1998	105	Clinton	R	HJRES121	Yes	
1999	106	Clinton	R	HJRES57	Yes	
2000	106	Clinton	R	HJRES103	Yes	
2001	107	G.W. Bush	R	HJRES50	Yes	
Granting PNTR to China:						
2000	106	Clinton	R	HR4444	Yes	

Table D.2—: Summary Statistics of Detrended Import Shocks at the Congressional District Level

	1990-2001		1997-2001		1993-1996		1990-1992	
	$S_{i,t+1}$	$S_{i,t}$	$S_{i,t+1}$	$S_{i,t}$	$S_{i,t+1}$	$S_{i,t}$	$S_{i,t+1}$	$S_{i,t}$
std	0.128	0.128	0.122	0.075	0.070	0.153	0.153	0.141
p5	-0.172	-0.175	-0.171	-0.106	-0.095	-0.186	-0.189	-0.187
p25	-0.077	-0.075	-0.084	-0.048	-0.043	-0.098	-0.093	-0.093
p50	-0.017	-0.016	-0.010	-0.009	-0.009	-0.020	-0.022	-0.022
p75	0.051	0.046	0.070	0.035	0.035	0.050	0.050	0.049
p95	0.239	0.231	0.214	0.144	0.139	0.316	0.293	0.272

Notes: In order to put the data on a comparable five-year scale, past import shocks over 1990-1992 are multiplied with the factor 5/2.

Additional results

E1. Evidence from congressional speech

The data on congressional speech is obtained from (Gentzkow, Shapiro and Taddy, 2019).[57] For each congressional district×year cell, we count the number of speeches that refer "China" together with a mention of trade issues or a mention of labor issues.[58] The number of relevant speeches is related to import shocks from China according to:

$$y_{it} = \sum_{c} \beta_c \mathbb{1}(t \in c) HighExposure_i + \gamma X_{it} + D_t + \varepsilon_{it}, \tag{E.1}$$

where y_{it} is the number of speeches related to the "China and trade" issue or the "China and labor" issue delivered by the representative of i in year t. Congressional districts are classified into groups based on their exposure to the increase in China's import penetration over the period 2001-2006. $HighExposure_i$ is an indicator variable equaling to 1 if i belongs to the top tercile of the exposure to import shock from China. We allow its effect to vary across different congressional sessions, and the differential effects are captured by coefficients β_c. X_i represents the total number of speeches delivered by the representative in district i and year t, and D_t denotes the year fixed effects. Standard errors are clustered at the congressional district level.

Panel A of Figure E.1 reports the estimates of β_c and the corresponding 90% confidence intervals. The estimated coefficients become positively significant from the 107th congress (2001-02) and on, when the China shock over 2001 to 2006 gradually revealed itself. In contrast, the estimates are smaller in magnitude and statistical insignificant over the earlier period 1989-1998. In panels B and C, we estimate equation (E.1) for Democrats and Republicans separately. For the purpose of comparison, y-axes in these two panels share the same scale. Two results emerge. First, after the China shock is realized, the representatives from the high-exposure districts raise the related trade and labor issues more often in their speeches, but such response is stronger for Democrats than Republicans. Second, it appears that Democrats started taking actions before the China shock is realized. Specifically, for Democrats the estimated β_c surged in 1999-2000 (i.e., the 106th congress), while for Republicans, the effect started picking up in 2001-2002. These patterns are consistent with the findings in section IV that (i) Democrats are more informed than Republicans about the China shock before it was fully realized, and (ii) Democratic legislators place more weights on the subconstituencies that would be adversely impacted by the future import penetration from

[57] See Gentzkow, Shapiro and Taddy (2018). We use the bound edition that collects the content for an entire Congress.

[58] We identify the issues that each speech cover by keywords. The keywords for trade issues include trade, export/exports and import/imports. The keywords for labor issues include labor, employment, unemployment, and job/jobs.

China.

E2. Pivotal voting and import shock including "NTR gap"

Our baseline empirical model of voting falls in the class of "expressive voting" models, where politicians do not incorporate the likelihood of the pivotality of their vote choice on the passage of the entire bill nor voters punish or reward politicians for the passage or failure of the bill at the polling station. In an expressive voting environment politicians have preferences over their individual choices and voters reward politicians for their individual support or opposition to a bill.

There are two main reasons for this modeling choice. First, it is empirically accurate. Politicians routinely campaign on their individual vote choices and attack each other based on the respective individual voting records, rather than on the outcome of a roll call vote. For example, in his 2020 primary campaign Senator Bernie Sanders remarked "*In the House and Senate, I voted against all of these terrible trade agreements, NAFTA, CAFTA, permanent normal trades relations with China.*" Former President Trump frequently attacked his challenger, Hillary Clinton, based on the her vote in the Senate in support of the war in Iraq during the 2016 presidential race.

Second, it is also a realistic assumption for decision making. Besides adding a layer of theoretical complexity in modeling decisions, pivotality concerns should be quantitatively relevant only for small deliberative bodies (e.g. the U.S. Supreme Court) where likelihood of pivotality is non trivial. Introducing pivotality concerns is a less appealing assumption when the set of agents has large cardinality (e.g. in the House of Representatives, that we consider). In these deliberative bodies the probability of actually being pivotal is close to zero and voters therefore should regard this component as quantitatively irrelevant. For instance, none of the votes on NTR bills were decided by a single vote.

That being said, it is interesting to engage with an analysis that incorporates pivotality and re-formulate the empirical analysis accordingly. We consider a simple pivotal voting model in which the deterministic component of the electoral support is determined by

$$\begin{aligned} h_t(d_{i,t}, S_{i,t+1}) &= \varphi_t + \gamma_t S_{i,t+1}^{NTR} \times \mathbb{1}\{d_{i,t} = \text{vote for } x_t\} + \gamma_t S_{i,t+1}^{NNTR} \times \mathbb{1}\{d_{i,t} = \text{vote for } q_t\} \\ &= \varphi_t + \gamma_t S_{i,t+1}^{NNTR} + \gamma_t (S_{i,t+1}^{NTR} - S_{i,t+1}^{NNTR}) \times \mathbb{1}\{d_{i,t} = \text{vote for } x_t\} \end{aligned}$$

where $\gamma_t < 0$ which captures the effect of realized shock on electoral support. The realized shock depends on the vote casted by i (i.e., the vote is pivotal). $S_{i,t+1}^{NTR}$ denotes the exposure to the import competition from China when NTR ends up being retained, $S_{i,t+1}^{NNTR}$ denotes the exposure to the import competition from China when NTR is revoked, and $S_{i,t+1}^{NTR} > S_{i,t}^{NNTR}$. Other elements of the model remain the same as the benchmark case in section I.

Therefore, the probability of voting in favor of China in this framework becomes

$$\Pr(Y_{i,t}=1|\mathcal{I}_{i,t}) = \Phi\left(\begin{array}{c} -\frac{1}{2}\left((x_t-\theta_i)^2-(q_t-\theta_i)^2\right) \\ +\tilde{\delta}\left(\mathbb{E}\left[V_{i,t+1}|x_t,\mathcal{I}_{i,t}\right]-\mathbb{E}\left[V_{i,t+1}|q_t,\mathcal{I}_{i,t}\right]\right)\end{array}\right),$$

where $\mathbb{E}\left[V_{i,t+1}|x_t,\mathcal{I}_{i,t}\right]-\mathbb{E}\left[V_{i,t+1}|q_t,\mathcal{I}_{i,t}\right] = \gamma_t\mathbb{E}\left[S^{NTR}_{i,t+1}-S^{NNTR}_{i,t+1}|\mathcal{I}_{i,t}\right]$. For the estimation, we use $S_{i,t+1}=S^{NTR}_{i,t+1}-S^{NNTR}_{i,t+1}$ to denote the differential China shocks under different trade regimes for the congressional district i, which is an aggregate of the differential trade shocks across commuting zones, $S_{j,t+1}=S^{NTR}_{j,t+1}-S^{NNTR}_{j,t+1}$ (where j denotes a CZ).

We can rationalize the formulation of $S_{j,t+1}$ based on a structural gravity equation. Specifically, the imports from China (c) by the US (u) in industry k is given by:

$$M_{cuk}=\frac{z_{ck}(1+\tau_{cuk})^{-\rho}}{\Phi_{uk}}M_{uk},$$

where z_{ck} is the composite of the state of technology, input cost and iceberg trade cost, which reflects China's supply capacity in industry k; τ_{cuk} measures the tariff imposed by the US on the imports of good k from China; ρ is the trade elasticity; M_{cuk} and M_{uk} denote, respectively, the imports by the US from China and the domestic absorption in the US. The term $\Phi_{uk}=\sum_i z_{ik}(1+\tau_{iuk})^{-\rho}$ captures the sourcing capacity of the US. Since China is a relatively small trade partner to the US in the baseline period (i.e., in the early the 1990s), for simplicity, we assume that the shocks to z_{ck} and τ_{cuk} have small impacts on Φ_{uk}. The formulation of the industry-level China shock in Autor, Dorn and Hanson (2013) given a tariff level is:

$$S_k=\frac{\Delta M_{cuk}}{M_{uk}}=\frac{\Delta z_{ck}\cdot(1+\tau_{cuk})^{-\rho}}{\Phi_{uk}},$$

where the magnitude of the shock hence depends on $(1+\tau_{cuk})$. Suppose there are two levels of tariffs τ^{NNTR}_{cuk} and τ^{NTR}_{cuk}, and $\tau^{NNTR}_{cuk}>\tau^{NTR}_{cuk}$. The corresponding shocks are denoted by S^{NNTR}_k and S^{NTR}_k. Hence,

$$\text{(E.2)}\qquad \begin{aligned} S^{NTR}_k-S^{NNTR}_k &= S^{NTR}_k\times\left[1-\left(\frac{1+\tau^{NNTR}_{cuk}}{1+\tau^{NTR}_{cuk}}\right)^{-\rho}\right] \\ &\approx \rho S^{NTR}_k\times\left(\ln(1+\tau^{NNTR}_{cuk})-\ln(1+\tau^{NTR}_{cuk})\right),\end{aligned}$$

where the approximation follows from the fact that $\ln(1-x)\approx -x$ when x has a small and positive value.[59] Note that S^{NTR}_k is the observed China supply shock. It is proxied by the surge in import penetration from China in eight other

[59]In our setting, $x=1-\left(\frac{1+\tau^{NNTR}_{cuk}}{1+\tau^{NTR}_{cuk}}\right)^{-\rho}$.

developed countries in Autor, Dorn and Hanson (2013), i.e., $(\frac{\Delta M^{oth}_{k,t+1}}{Y_{k,t}+M_{k,t}-X_{k,t}})$.

Motivated by the industry-level shocks as in equation (E.2), we construct the China shock at the CZ level that embeds different policy regimes according to

$$S_{j,t+1} = \sum_k \frac{L_{jk,t}}{L_{j,t}} \left(NTRGap_{k,t} \times \frac{\Delta M^{oth}_{k,t+1}}{Y_{k,t} + M_{k,t} - X_{k,t}} \right), \tag{E.3}$$

where $NTRGap_{k,t} = \ln(1 + NonNTR\,Tariff_{k,t}) - \ln(1 + NTR\,Tariff_{k,t})$. Note that this measure drops the parameter ρ, and it is proportional to its theoretical counterpart: $\sum_k \frac{L_{jk,t}}{L_{j,t}}(S^{NTR}_{k,t+1} - S^{NNTR}_{k,t+1})$. The term $NTRGap_{k,t} \times \frac{\Delta M^{oth}_{k,t+1}}{Y_{k,t}+M_{k,t}-X_{k,t}}$ is proportional to the potential reduction in the China shock at the industry level if imports from China had faced the non-Most Favored Nation tariffs. Therefore, *ceteris paribus*, legislators are less likely to vote in favor of China if the local economy specializes more in industries that face larger supply shocks from China, and will receive more tariff protection if China's NTR status is revoked. The past shock is constructed analogously according to:

$$S_{j,t} = \sum_k \frac{L_{jk,t-5}}{L_{j,t-5}} \Big(NTRGap_{k,t} \times \frac{\Delta M^{oth}_{k,t}}{Y_{k,t-5} + M_{k,t-5} - X_{k,t-5}} \Big). \tag{E.4}$$

These CZ level measures are then mapped to the CD level measures ($S_{i,t+1}$ and $S_{i,t}$) based on the procedure described in the main text. The summary statistics are reported in Table E.5.

Table E.6 repeats our baseline analysis using the alternative measures (E.3) and (E.4). As it can be seen, the results remains similar to our main estimates. On top of the considerations above, therefore, pivotality considerations appear not quantitatively damning in our case. Note that the estimate of δ is larger in magnitude in this Table, but it is simply due to the fact that the standard deviation of measure (E.3) is about one third of the baseline measure.

E3. Introducing export shocks

Similar to Feenstra, Ma and Xu (2019), we construct the future export shock from China in CZ j according to:

$$S^X_{j,t+1} = \sum_k \frac{L_{jk,t}}{L_{j,t}} \frac{\Delta X^{oth}_{k,t+1}}{Y_{k,t}}.$$

As with import shocks, we use the change in export of good k to China from eight other (non-U.S.) high-income countries over 5 years in the future to capture the shift in export demand from China. The past export shock is constructed accordingly. These the CZ level measures are then mapped to the CD level based on the procedure in the main text.

In Table E.7, we augment the baseline model with export shocks. In the specification, we assume $Z_{i,t} = \{S_{i,t}, S^X_{i,t}, ShareMfg_{i,t}, \theta_i\}$. We find a positive effect of export opportunities on the probability of a vote in favor of China's NTR status. Yet, the information tests and the other estimates remain largely consistent with the main findings.

Robustness to potential model misspecifications

F1. Robustness to violations of the rational expectation assumption

Consider that the data generating process of future shock from China is

$$S_{i,t+1} = \beta_0 + \beta_1 ShareMfg_{i,t} + \beta_2 S_{i,t} + \varepsilon_{i,t+1}, \tag{F.1}$$

where $\mathbb{E}(\varepsilon_{i,t+1}|ShareMfg_{i,t}, S_{i,t}) = 0$. In this exercise, we assume that $\mathcal{I}_{i,t} = \{\theta_i, ShareMfg_{i,t}, S_{i,t}\}$.

Case I: Expectational errors correlated with $S_{i,t}$

Now, suppose that politicians form expectation in the following way:

$$\mathbb{E}(S_{i,t+1}|\mathcal{I}_{i,t}) = \beta_0 + \beta_1 ShareMfg_{i,t} + (\beta_2 + \rho_s)S_{i,t}. \tag{F.2}$$

When $\rho_s < 0$, it implies that politicians in the regions with a positive (respectively, negative) past shock under-predict (respectively, over-predict) the future shock. The opposite is the case when $\rho_s > 0$. The expectational error is then:

$$\nu_{i,t+1} = S_{i,t+1} - \mathbb{E}(S_{i,t+1}|\mathcal{I}_{i,t}) = \varepsilon_{i,t+1} - \rho_s S_{i,t}.$$

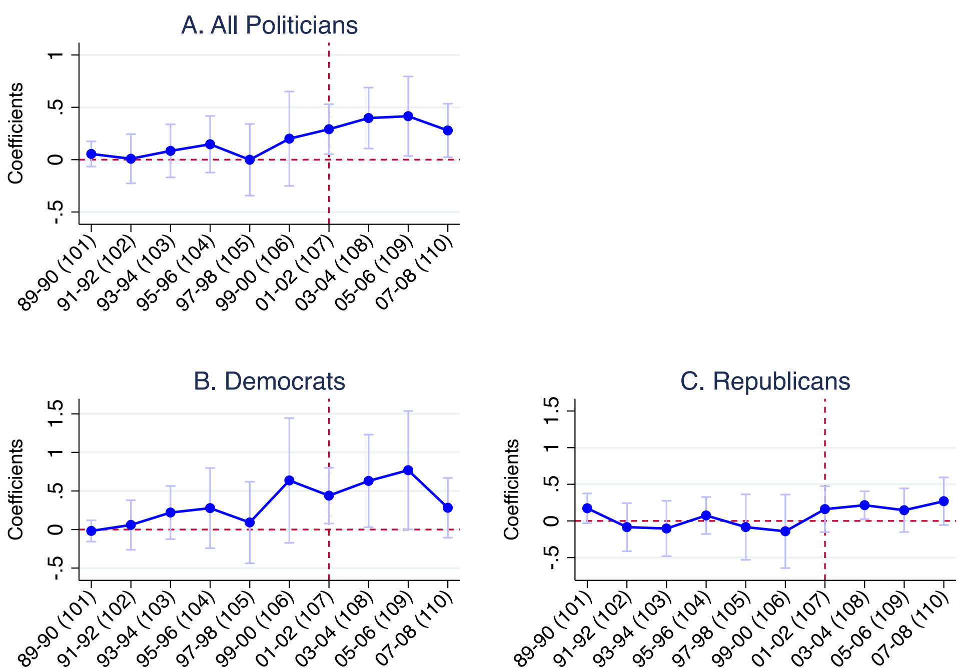
In such a scenario, the rational expectation assumption is violated, i.e., $\mathbb{E}(\nu_{i,t+1}|\mathcal{I}_{i,t}) \neq 0$.

We infer the potential bias when the rational expectation assumption is violated by conducting a Monte Carlo simulation as follows. First, we simulate the data on voting outcomes based on the expected value in equation (F.2). Second, we then naively estimate our baseline model. To simulate $S_{i,t+1}$, we set $\beta_0 = 0$, $\beta_1 = 0.721$, $\beta_2 = 0.184$. $\varepsilon_{i,t+1}$ are drawn from the normal distribution with mean 0 and standard deviation of $\sigma_\varepsilon = 0.525$. We take $S_{i,t}$, $ShareMfg_{i,t}$ and θ_i from the real data over 1998-2001, and simulate different cases with $\rho_s = \{-0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.8\}$. The deviation from the rational expectation assumption is more severe the larger the absolute value of ρ_s is. Figure F.1 presents the estimated confidence sets of the parameters a, b and δ corresponding to different ρ_s under $Z_{i,t} = \mathcal{I}_{i,t}$.

There are four main findings. First, when $\rho_s = 0$ (i.e., the rational expectation assumption holds), the 95% CS always contain the true parameters. Second, when $\rho_s > 0$, we find that (i) confidence sets get wider for all parameters, and (ii) for δ, the CS shifts downward, and may not cover the true value of δ when ρ_s is large enough. The intuition for (i) can be illustrated by the m_l^{ob} inequality. We start from the identity function

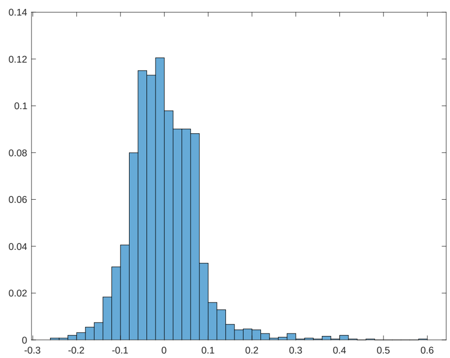
$$\mathbb{E}\left[(1 - Y_{i,t})\frac{\Phi(a_t\theta_i + b_t + \delta_t S_{i,t+1} - \delta_t \nu_{i,t+1})}{1 - \Phi(a_t\theta_i + b_t + \delta_t S_{i,t+1} - \delta_t \nu_{i,t+1})} - Y_{i,t}\middle|\mathcal{I}_{i,t}\right] = 0. \tag{F.3}$$

Figure E.1. : Number of Speeches with Mentions of China & Labor or China & Trade Issues

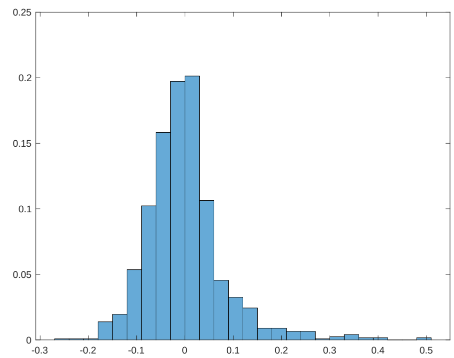


Notes: This figure plots the estimated coefficients β_c in equation (E.1). Error bands show 90% confidence intervals. Standard errors are clustered at the congressional district level. The vertical red dashed line indicates the time of China's accession to the WTO.

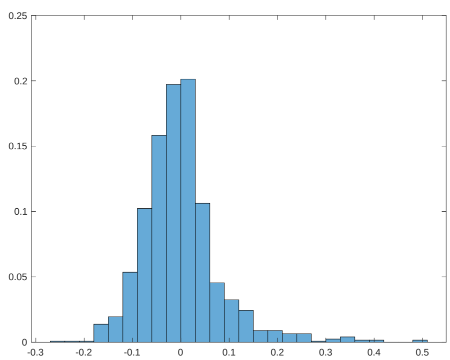
Figure E.2. : Distribution of Expectational Errors



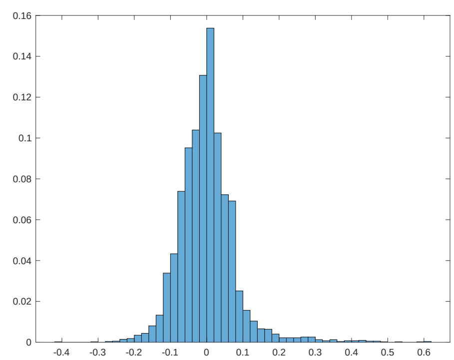
A. 1997-2001



B. 1993-1996



C. 1990-1992



D. 1990-2001

Notes: The figure plots the distribution of expectational errors based on the Baseline information set, i.e., $\varepsilon_{i,t+1} = S_{i,t+1} - \mathbb{E}[S_{i,t+1}|\mathcal{I}^b_{i,t}]$. Specifically, the expectational errors are the residuals from the OLS regression: $S_{i,t+1} = \beta_0 + \beta_1\theta_i + \beta_2 ShareMfg_{i,t} + \beta_3 S_{i,t} + \varepsilon_{i,t+1}$.

Table E.1—: Parameter Confidence Sets and Specification Test p-values: Heterogeneity by Party (Minimal Information)

Period	Group	CS of a	CS of b	CS of δ	p-value BP	p-value RC	p-value RS	Num obs.
1997-2001	Democracts	[-0.475, 5.300]	[-0.125, 1.938]	[-9.742, -0.730]	0.990	0.990	0.990	1229
	Republicans	[-1.950, -1.125]	[0.912, 1.288]	[-1.300, 0.013]	0.125	0.110	0.110	1326
1993-1996	Democracts	[0.875, 6.300]	[0.950, 2.975]	[-17.000, 0.062]	0.990	0.990	0.990	888
	Republicans	[-1.312, -0.562]	[0.800, 1.075]	[-0.200, 2.800]	0.070	0.050	0.050	806
1990-1992	Democracts	[0.625, 3.500]	[-0.200, 0.600]	[-2.500, 6.000]	0.665	0.665	0.665	745
	Republicans	[-1.200, 1.500]	[-0.350, 0.550]	[-1.100, 3.700]	0.465	0.465	0.465	485
1990-2001	Democracts	[0.400, 3.300]	[0.290, 1.225]	[-8.565, -0.517]	0.910	0.885	0.885	2862
	Republicans	–	–	–	0.010	0.010	0.010	2617

Notes: The estimation in this table is based on the Minimal information $Z_{i,t} = \{ShareMfg_{i,t}.\theta_i\}$.

Setting aside the problem that $v_{i,t+1}$ and $S_{i,t+1}$ are correlated for now, a larger ρ_s increases the variation of $\nu_{i,t+1}$, which mechanically raises the left-hand-side of the following inequality and hence reduces the obtained lower bounds:

$$\mathbb{E}\left[(1-Y_{i,t})\frac{\Phi(a_t\theta_i+b_t+\delta_t S_{i,t+1})}{1-\Phi(a_t\theta_i+b_t+\delta_t S_{i,t+1})}-Y_{i,t}\middle|\mathcal{I}_{i,t}\right]>0.$$

A similar argument applies to inequalities m_u^{ob}, m_l^{rp} and m_u^{rp}. To glean some intuitions for (ii), we note that when $\rho_s > 0$, $v_{i,t+1}$ and $S_{i,t+1}$ are negatively correlated. When we replace $\mathbb{E}(S_{i,t+1}|\mathcal{I}_{i,t})$ by $S_{i,t+1}$, the standard omitted variable bias problem drives the estimate of δ downward.

Third, when $\rho_s < 0$, we find that (i) confidence set of δ shits upward and may not cover the true value of δ when ρ_s is negative enough, and (ii) the specification is rejected (i.e., the CS is empty) when $\rho_s = -0.8$. On (i), the bias is induced by the similar omitted variable problem discussed above. On (ii), it is because we set $\beta_2 = 0.721$, and $\beta_2 + \rho_s \approx 0$ when $\rho_s = -0.8$. Note that in such a case, $S_{i,t}$ takes little role in forming expectation on $S_{i,t+1}$ (see equation F.2), and hence it is as if politicians have no information on $S_{i,t}$. (More precisely, it is as if politicians only have information on $\{\theta_i, ShareMfg_{i,t}\}$, while researchers mistakenly assume $Z_{i,t} = \{\theta_i, ShareMfg_{i,t}, S_{i,t}\}$. Hence, the specification is rejected.)

Fourth, even when $\rho_s \neq 0$, the confidence sets of a and b always contain the true values. This is because in our data, the correlation between θ_i and $S_{i,t+1}$ is only 0.06. Therefore, an error correlated with $S_{i,t+1}$ is unlikely to lead to a severe bias in the estimated coefficient for θ_i.

How likely are our baseline estimates of δ to be severely biased due to a violation of the rational expectation assumption? Based on the simulated results in Figure F.1, the potential biases are more concerning when $|\rho_s| > 0.4$. We also note that the standard deviation of $\varepsilon_{i,t+1}$ in equation (F.1) is 0.525 (which is set to match the empirical distribution of the regression residuals), while the standard deviation of $S_{i,t}$ is around 1. This implies that the standard deviation of $\rho_s S_{i,t}$ is approximately ρ_s. In other words, the biases are more severe when the variation of irrational expectational errors is at least as large as the variation of future China shock that is unexplained by observables $ShareMfg_{i,t}$ and $S_{i,t}$.

We now turn the attention to the case $Z_{i,t} \not\subseteq \mathcal{I}_{i,t}$. Specifically, we use the simulated data with different values of ρ_s, and estimate the model based on the assumption of perfect foresight. We always reject the case of perfect foresight. In sum, it does not seem that the violation of rational expectation assumption hinders our ability to reject the model with misspecified information set.

Case II: Expectational errors correlated with θ_i

We next consider the scenario in which the expectational errors are systematically correlated with ideology. Specifically, the expectation of future China shock is given by

$$\mathbb{E}(S_{i,t+1}|\mathcal{I}_{i,t}) = \beta_0 + \beta_1 ShareMfg_{i,t} + \beta_2 S_{i,t} + \rho_\theta \theta_i. \tag{F.4}$$

The corresponding expectational error is

$$\nu_{i,t+1} = S_{i,t+1} - \mathbb{E}(S_{i,t+1}|\mathcal{I}_{i,t}) = \varepsilon_{i,t+1} - \rho_\theta \theta_i.$$

When $\rho_\theta < 0$, politicians with a positive (respectively, negative) θ under-predict (respectively, over-predict) the future shock. This case implies that Republicans are more likely to under-predict the future shock than Democrats. The opposite is the case when $\rho_\theta > 0$. In such scenarios, the rational expectation assumption is again violated, i.e., $\mathbb{E}(\nu_{i,t+1}|\mathcal{I}_{i,t}) \neq 0$.

Based on the same parameterization of $\{a, b, \delta, \sigma_\varepsilon, \beta_0, \beta_1, \beta_2\}$ as in section F.F1, we simulate data for different cases with $\rho_\theta = \{-1.6, -1.2, -0.8, -0.4, 0, 0.4, 0.8, 1.2, 1.6\}$. The deviation from the rational expectation assumption is more severe the larger the absolute value of ρ_θ is. Figure F.2 shows the estimated confidence sets of the parameters a, b and δ corresponding to different ρ_θ under $Z_{i,t} = \mathcal{I}_{i,t}$.

Two patterns emerge. First, the confidence sets are wide when ρ_θ is further away from 0. Again, this is because the variation of $\nu_{i,t+1}$ is larger when $|\rho_\theta|$ is larger. Second, the CS of a shifts downward (respectively, upward) when ρ_θ becomes more positive (respectively, negative). Specifically, when $\rho_\theta > 0$, $v_{i,t+1} = \varepsilon_{i,t+1} - \rho_\theta \theta_i$ and θ_i are negatively correlated. When we replace $\mathbb{E}(S_{i,t+1}|\mathcal{I}_{i,t})$ by $S_{i,t+1}$ in the estimating equation, the standard omitted variable bias problem drives the estimate of a downward. The opposite is the case when $\rho_\theta < 0$.

Importantly, even when $\rho_\theta \neq 0$, the confidence sets of b and δ always contain the true values. Moreover, for the range of ρ_θ under consideration, we never reject the specification that politicians have information on $Z_{i,t} = \{\theta_i, ShareMfg_{i,t}, S_{i,t}\}$. This is because in our data, the correlation between θ_i and $S_{i,t+1}$ is statistically zero, and the potential omitted variable correlated with θ_i has little impact on estimating δ and inferring the information set. How severe is the bias associated with the estimate of a? Based on the simulated results in Figure F.2, the potential biases are more concerning when $|\rho_\theta| \geq 1.2$. In our data, the standard deviation of θ_i is 0.412. When $|\rho_\theta| = 1.2$, the standard deviation of $\rho_\theta \theta_i$ is 0.494, which is almost as large as the total variation of future China shock that is unexplained by observables $ShareMfg_{i,t}$ and $S_{i,t}$. Therefore, with the presence of irrational expectational errors that are moderately correlated with ideology, the baseline estimated CS of a is unlikely to be severely biased.

For a wide range of ρ_θ, we always reject the case of perfect foresight. Hence, the violation of rational expectation assumption due to the correlation between expectational errors and ideology does not seem to affect the power of the specification tests that we employ.

Case III: Expectation error is a Non-Zero Constant

We now consider the case when all politicians under-predict (or over-predict) the China shock simultaneously. To probe this question, we simulate data assuming that the expectation on future China shock conditional on $\mathcal{I}_{i,t}$ takes the following form:

$$\mathbb{E}(S_{i,t+1}|\mathcal{I}_{i,t}) = \beta_0 + \rho_0 + \beta_1 ShareMfg_{i,t} + \beta_2 S_{i,t}. \tag{F.5}$$

Hence, the expectational error in this case is

$$\nu_{i,t+1} = S_{i,t+1} - \mathbb{E}(S_{i,t+1}|\mathcal{I}_{i,t}) = \varepsilon_{i,t+1} - \rho_0.$$

The rational expectation assumption is again violated, i.e., $\mathbb{E}(\nu_{i,t+1}|\mathcal{I}_{i,t}) \neq 0$.

Figure F.3 presents the confidence sets for the simulated data for different cases with $\rho_0 = \{-0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.8\}$. The simulation is based on the same parameterization of $\{a, b, \delta, \sigma_\varepsilon, \beta_0, \beta_1, \beta_2\}$ as in section F.F1. Again, the absolute value of ρ_0 governs the extent to which the rational expectation assumption is violated. With the specification $Z_{i,t} = \mathcal{I}_{i,t}$, we find that (i) the confidence sets are wider the larger is $|\rho_0|$, and (ii) the confidence set of b shifts when $\rho_0 \neq 0$, and does not cover the true value of b when $|\rho_0|$ gets larger.

Importantly, even when $\rho_0 \neq 0$, the confidence sets of a and δ always contain the true values. Intuitively, this is analogous to the scenario that the OLS estimates other than the constant coefficient remain consistent when an explanatory variable is changed by a constant. Since the main parameters of interest in our analysis are a and δ, we consider the case with $\rho_0 \neq 0$ less concerning. As with aforementioned cases, we always reject the specification with perfect foresight. Hence, the violation of rational expectation assumption due to a constant expectational error does not seem to affect the power of the specification test.

F2. Ideology-dependent information sets

In this subsection, we simulate the data where politicians with $\theta \leq 0$ have access to information on $\mathcal{I}^b_{i,t} = \{\theta_i, ShareMfg_{i,t}, S_{i,t}\}$, while a share ($\pi$) of politicians with $\theta > 0$ only have information on $\mathcal{I}^m_{i,t} = \{\theta_i, ShareMfg_{i,t}\}$. Their expectations on the future China shock are rational, and are given respectively by:

$$\mathbb{E}(S_{i,t+1}|\mathcal{I}^b_{i,t}) = \beta_0 + \beta_1 ShareMfg_{i,t} + \beta_2 S_{i,t},$$

and

$$\mathbb{E}(S_{i,t+1}|\mathcal{I}^m_{i,t}) = \beta_0 + \beta_1 ShareMfg_{i,t}.$$

We simulate different cases with $\pi = \{0.5, 0.66, 0.75, 0.9, 1\}$. We pool the data and test the specification with baseline information set $\mathcal{I}^b_{i,t}$. Hence, the model is misspecified for some individuals. Specifically, the case $\pi = 1$ is to capture the scenario that none of the Republicans have the baseline information.

The results are reported in Figure F.4. We find that for the extreme case $\pi = 1$, the 95% confidence set of δ is biased towards 0. The corresponding p-value of the RC test is 0.065. For more realistic values of π (i.e, 0.5, 0.66 and 0.75), the CS of δ still covers the true value. In sum, the estimation based on moment inequalities is largely robust to the case where the information set is partially misspecified.

Table E.2—: Parameter Estimates and Content of Information Sets
(Extended Model: Including Campaign Contributions from Interest Groups)

Period	CS of a (Ideology)	CS of b (Constant)	CS of δ (Import Shock)	CS of χ (Money)	p-value BP	p-value RC	p-value RS	Num obs.
Panel A: Including Campaign Contributions from Business								
1997-2001	[-0.340, 0.940]	[0.233, 0.533]	[-4.867, -0.067]	[0.133, 0.333]	0.845	0.845	0.845	2546
1993-1996	[-0.833, 0.000]	[0.680, 0.840]	[-5.533, 0.533]	[0.067, 0.233]	0.265	0.265	0.265	1688
1990-1992	[0.793, 1.527]	[-0.233, -0.133]	[-1.440, -0.040]	[-0.067, 0.133]	0.130	0.130	0.130	1229
1990-2001	[0.000, 0.600]	[0.200, 0.367]	[-4.200, -0.200]	[0.100, 0.267]	0.950	0.950	0.950	5463
Panel B: Including Campaign Contributions from Labor								
1997-2001	[-0.220, 0.480]	[0.180, 0.260]	[-2.200, -0.067]	[-0.127, -0.040]	0.125	0.125	0.125	2546
1993-1996	[-0.827, -0.093]	[0.600, 0.720]	[-3.467, -0.533]	[-0.127, -0.040]	0.200	0.200	0.200	1688
1990-1992	[0.000, 1.400]	[-0.367, -0.167]	[-2.467, -0.067]	[-0.200, 0.000]	0.185	0.185	0.185	1229
1990-2001	[-0.140, 0.660]	[0.220, 0.340]	[-2.400, -0.267]	[-0.147, 0.013]	0.845	0.845	0.845	5463

Notes: The estimation in this table is based on the assumption that the information set possessed by politicians contains $S_{i,t}$, $ShareMfg_{i,t}$, θ_i, and the campaign contributions from the corresponding interest groups during the congressional cycle.

Table E.3—: Excluding Politicians Working in the Committee on Commerce, and the Committee on Ways and Means

Period	CS of a	CS of b	CS of δ	p-value BP	p-value RC	p-value RS	Num obs.
1997-2001	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[0.500, 0.920]	[0.125, 0.237]	[-2.550, -0.075]	0.745	0.745	0.745	2012
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[0.440, 0.930]	[0.125, 0.275]	[-2.400, -0.150]	0.815	0.815	0.815	2012
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1}\|S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	–	–	–	0.010	0.010	0.010	2012
1993-1996	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[-0.425, 0.075]	[0.575, 0.700]	[-3.163, -0.012]	0.315	0.315	0.315	1361
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[-0.230, -0.130]	[0.600, 0.630]	[-1.650, -0.700]	0.055	0.050	0.050	1361
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1}\|S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	–	–	–	0.020	0.020	0.020	1361
1990-1992	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[0.750, 1.600]	[-0.425, -0.175]	[-1.250, 2.375]	0.955	0.955	0.955	1001
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[0.988, 1.512]	[-0.350, -0.200]	[-1.458, 0.093]	0.225	0.220	0.220	1001
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1}\|S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	[0.950, 1.625]	[-0.350, -0.188]	[-1.690, 0.248]	0.365	0.365	0.365	1001
1990-2001	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[0.487, 0.638]	[0.168, 0.220]	[-1.240, -0.290]	0.120	0.120	0.120	4374
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[0.463, 0.675]	[0.183, 0.250]	[-1.525, -0.480]	0.205	0.205	0.205	4374
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1}\|S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	–	–	–	0.010	0.010	0.010	4374

Notes: For the case of perfect foresight, we assume that in addition to $S_{i,t}$, $ShareMfg_{i,t}$ and θ_i, politicians also possess information that is orthogonal to these covariates, i.e. $S_{i,t+1} - E[S_{i,t+1}|S_{i,t}, ShareMfg_{i,t}, \theta_i]$.

Table E.4—: Parameter Estimates and Content of Information Sets: Congressional Voting on Annual Foreign Operations Appropriations Bill

Period	CS of a	CS of b	CS of δ	p-value of BP	p-value RC	p-value RS	Num obs.
1997-2001	[-0.825, -0.275]	[0.470, 0.600]	[-1.363, 0.300]	0.520	0.500	0.500	2068
1993-1996	[-0.650, -0.350]	[0.812, 0.900]	[-0.350, 1.675]	0.085	0.085	0.085	1685
1990-1992	[-0.900, -0.275]	[0.490, 0.670]	[-0.475, 1.250]	0.370	0.370	0.370	1241
1990-2001	[-0.575, -0.425]	[0.620, 0.657]	[-0.240, 0.248]	0.080	0.080	0.080	5003

Notes: The estimation in this table is based on the Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_{i,t}\}$. The bills included are: HR5114 (1990), HR2621 (1991), HR5368 (1992), HR2295 (1993), HR4426 (1994), HR1868 (1995), HR3540 (1996), HR2159 (1997), HR4569 (1998), HR2606 (1999), HR4811 (2000), and HR2506 (2001).

Table E.5—: Summary Statistics of Detrended "NTR gap" Import Shocks at the Congressional District Level

	1990-2001		1997-2001		1993-1996		1990-1992	
	$S_{i,t+1}$	$S_{i,t}$	$S_{i,t+1}$	$S_{i,t}$	$S_{i,t+1}$	$S_{i,t}$	$S_{i,t+1}$	$S_{i,t}$
std	0.038	0.043	0.039	0.027	0.023	0.054	0.052	0.050
p5	-0.050	-0.057	-0.052	-0.038	-0.032	-0.063	-0.061	-0.062
p25	-0.022	-0.025	-0.025	-0.017	-0.014	-0.035	-0.031	-0.032
p50	-0.004	-0.005	-0.004	-0.003	-0.003	-0.008	-0.007	-0.007
p75	0.013	0.014	0.019	0.011	0.010	0.016	0.012	0.019
p95	0.066	0.080	0.072	0.052	0.045	0.111	0.105	0.094

Notes: In order to put the data on a comparable five-year scale, past import shocks over 1990-1992 are multiplied with the factor 5/2.

Table E.6—: Parameter Confidence Sets and Specification Test p-values with "NTR gap" Import Shocks

Period	CS of a	CS of b	CS of δ	p-value BP	p-value RC	p-value RS	Num obs.
1997-2001	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[0.260, 0.740]	[0.060, 0.270]	[-9.900, -1.100]	0.990	0.990	0.990	2564
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[0.450, 0.750]	[0.185, 0.285]	[-5.225, -0.730]	0.285	0.285	0.285	2564
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1} \| S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	–	–	–	0.010	0.010	0.010	2564
1993-1996	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[-0.310, 0.140]	[0.507, 0.688]	[-10.875, 1.875]	0.270	0.270	0.270	1698
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[-0.310, 0.110]	[0.613, 0.733]	[-9.500, -1.500]	0.215	0.215	0.215	1698
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1} \| S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	–	–	–	0.010	0.010	0.010	1698
1990-1992	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[0.650, 1.587]	[-0.420, -0.035]	[-6.700, 5.000]	0.990	0.990	0.990	1232
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[1.062, 1.400]	[-0.260, -0.170]	[-4.375, -0.875]	0.095	0.080	0.080	1232
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1} \| S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	[1.062, 1.475]	[-0.280, -0.170]	[-3.875, -0.500]	0.130	0.125	0.125	1232
1990-2001	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[0.330, 0.585]	[0.140, 0.245]	[-5.250, -0.750]	0.625	0.625	0.625	5494
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[0.495, 0.630]	[0.240, 0.280]	[-4.550, -2.188]	0.125	0.125	0.125	5494
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1} \| S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	–	–	–	0.010	0.010	0.010	5494

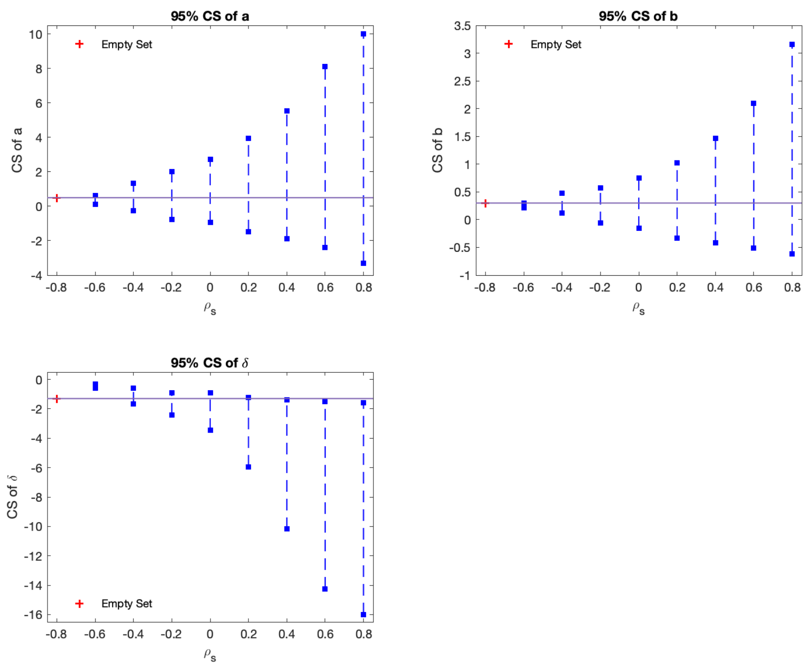
Notes: For the case of perfect foresight, we assume that in addition to $S_{i,t}$, $ShareMfg_{i,t}$ and θ_i, politicians also possess information that is orthogonal to these covariates, i.e. $S_{i,t+1} - E[S_{i,t+1}|S_{i,t}, ShareMfg_{i,t}, \theta_i]$.

Table E.7—: Parameter Estimates and Content of Information Sets: Model Including Export Shocks

Period	CS of a (Ideology)	CS of b (Constant)	CS of δ (Import Shock)	CS of χ (Export Shock)	p-value BP	p-value RC	p-value RS	Num obs.
1997-2001	[0.320, 1.120]	[0.120, 0.320]	[-5.600, -2.333]	[8.947, 15.100]	0.430	0.430	0.430	2564
1993-1996	[-0.367, 0.233]	[0.607, 0.767]	[-4.667, -0.333]	[9.480, 16.200]	0.290	0.280	0.280	1698
1990-1992	[0.800, 1.733]	[-0.333, -0.100]	[-3.267, 0.633]	[-7.500, 10.300]	0.645	0.645	0.645	1232
1990-2001	[0.320, 0.960]	[0.087, 0.367]	[-6.400, -2.133]	[11.000, 17.500]	0.750	0.750	0.750	5494

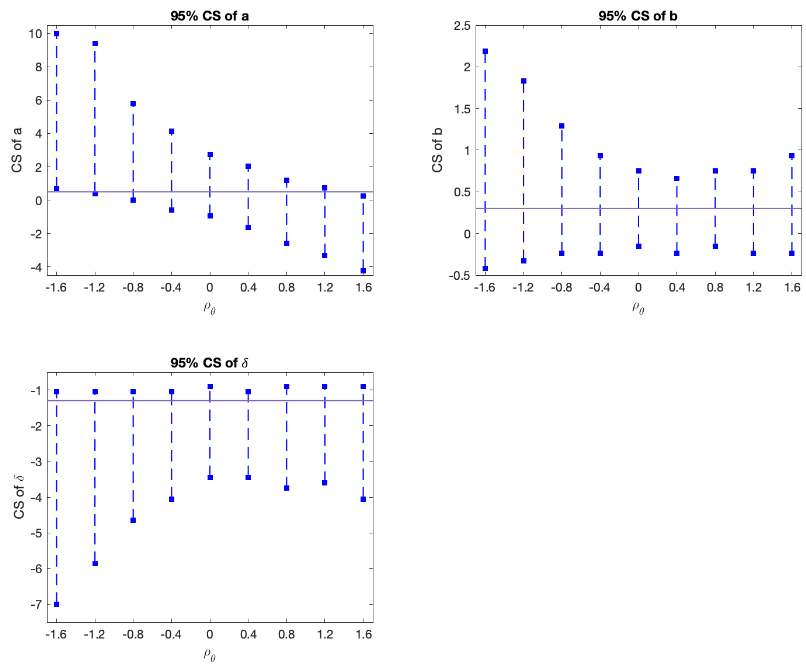
Notes: The estimation in this table is based on the assumption that the information set possessed by politicians contains $S_{i,t}$, $S^X_{i,t}$, $ShareMfg_{i,t}$, and θ_i.

Figure F.1. : Parameter Confidence Sets When the Rational Expectation Assumption is Violated
(Expectational Errors are Correlated with $S_{i,t}$)



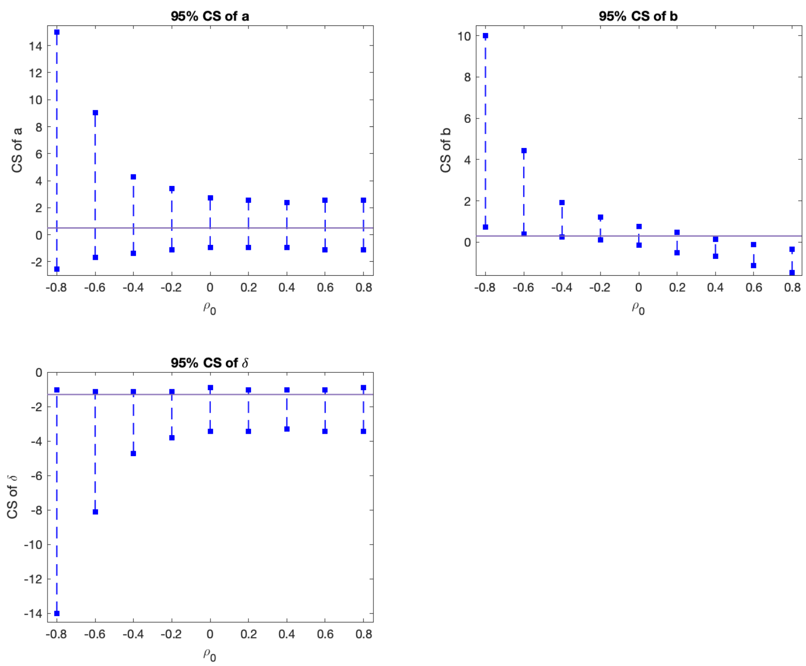
Note: The figure show the 95% confidence sets for the parameters for specifications with different ρ_s. The purple lines represent the true parameters of a, b, and δ, respectively.

Figure F.2. : Parameter Confidence Sets When the Rational Expectation Assumption is Violated
(Expectational Errors are Correlated with θ_i)



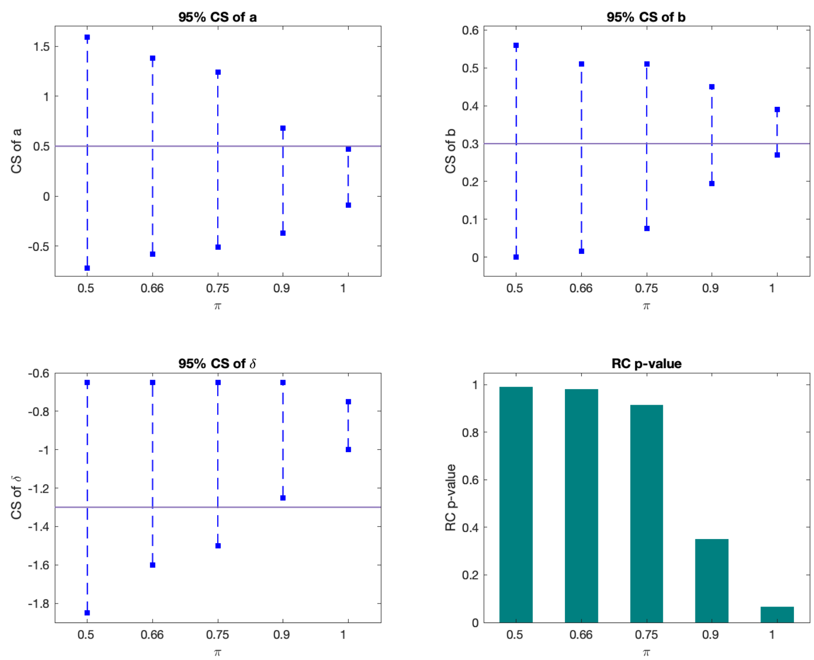
Note: The figure show the 95% confidence sets for the parameters for specifications with different ρ_θ. The purple lines represent the true parameters of a, b, and δ, respectively.

Figure F.3. : Parameter Confidence Sets When the Expectation Error is a Non-Zero Constant
(Expectational Errors is ρ_0)



Note: The figure show the 95% confidence sets for the parameters for specifications with different ρ_θ. The purple lines represent the true parameters of a, b, and δ, respectively.

Figure F.4. : Parameter Confidence Sets When the Information Set is Misspecified for Some Politicians with $\theta > 0$



Note: The figure show the 95% confidence sets for the parameters for specifications with different ρ_θ. The purple lines represent the true parameters of a, b, and δ, respectively.