Did U.S. Politicians Expect the China Shock?

Matilde Bombardini, Bingjing Li
and Francesco Trebbi*

March 2022

Abstract

Information sets, expectations, and preferences of politicians are fundamental, but unobserved determinants of their policy choices. Employing repeated votes in the U.S. House of Representatives on China's Normal Trade Relations status during the two decades straddling China's WTO accession, we apply a moment inequality approach designed to deliver consistent estimates under weak informational assumptions on the information sets of members of Congress. This methodology offers a robust way to test hypotheses about what information politicians have at the time of their decision and to estimate the weight that constituents, ideology, and other factors have in policy making and voting.

1   Introduction

The China shock, the large surge in imports from China that started in the 1990s and that has turned China into one of the U.S. main trading partners, has received broad attention in academia and policy making. Although early research by Autor et al. (2013, 2016) and Pierce and Schott (2016) focused mainly of its labor market effects, a large literature has expanded the analysis to health, social, and political consequences of the shock.[1] While in hindsight China's entry in the U.S. market may seem like a preordained outcome, in a series of roll call votes during the 1990s members of Congress were faced with the choice of allowing China to maintain its Normal Trade Relations (NTR), and ultimately obtain Permanent NTR, status.

In this paper we ask to what degree members of the U.S. House of Representatives were informed about the consequences of the China shock for their voters and how much the expected impact on their constituents affected their support in favor or against China's NTR status.[2] These two questions are intrinsically related and point to the difficulty of modeling forward-looking expectations and decisions by policy makers in a context where these choices depend on consequences that are not known at the time of the vote. To this goal, we present a model and an estimation strategy of the decisions and expectations of law makers. In this, we depart from the empirical political economy literature on legislative voting (Poole and Rosenthal, 1997; Heckman and Snyder, 1997; Clinton et al., 2004; Canen et al., 2020) and extend it in a distinct direction, as a formal analysis of law makers' information sets and expectations does not typically figure in standard empirical models of voting.

To provide intuition of why appropriate modeling of policy makers expectations is relevant, consider the following. A naïve approach to estimating the importance of constituent interests may be the replacement of the politician's expectations of the China shock with their realized values. However, assuming that politicians are perfectly informed about future shocks when they are not, necessarily implies a downward bias in the coefficient that measures the preference weight placed on constituent interests within a discrete choice voting model. This is due to an intuitive error-in-variables argument (i.e. the mechanical negative covariance between the expectational error and the realized future value of the shock). Without correction, a small coefficient may be interpreted as low responsiveness of politicians to subconstituents' fortunes, possibly indicating a political accountability problem (Kalt and Zupan, 1984, 1990). In reality, a small coefficient may be as well the result of assuming that politicians are better informed than they truly are. Yet, this rather points to a limited expertise or insufficient information acquisition of legislators (Krehbiel,

---

[1]Among the others, see Greenland and Lopresti (2016), Feler and Senses (2017), Autor et al. (2019), Greenland et al. (2019), Autor et al. (2020) and Pierce and Schott (2020).

[2]The role of electoral constituencies and subconstituencies in driving the behavior of members of Congress has played a central role in the analysis of policy support in Washington at least since Fenno (1978); Peltzman (1984). See also Mian et al. (2010, 2014) for more recent applications.

1992). Because these interpretations have distinct policy implications and they call for different remedies, it seems relevant to be able to distinguish between them.

To address the estimation challenge, we link the political economy literature to a separate strand of the international trade literature. The novel moment inequality methodology of Dickstein and Morales (2018), developed in the context of the decision of firms to export to foreign markets, allows us to consistently operate within an expectational environment where what belongs to the information set of the decision maker is only partially observed. That is, this moment inequality approach only requires the econometrician to know a subset of the information available to the politician at the time of his or her vote for the consistent estimation of the parameters – a much less demanding restriction. This approach turns out to be particularly informative in our context. It allows us to estimate a voting model under general assumptions about the politician's information set and expectations, and therefore to answer the question of how much politicians cared about the China shock in the first place.

Naturally, economic consequences on their constituents were not the only considerations affecting individual law makers' support for NTR and our model accommodates these features.[3] It is generally believed that several members of Congress voted to withhold NTR status in order to affect China's position on human rights, as the series of yearly roll call votes on China's NTR status between 1990 and 2001 started after the Tiananmen Square events of 1989. We therefore also allow the voting behavior to depend on the ideological position of the legislator together the expected electoral cost of supporting China's NTR status, taking into consideration the district-specific impact of China's continued and growing exports to the U.S.. Further, in the utility function, ideology also captures the position of the legislator towards free trade policy, that is the value the politician places on the collective gains from maintaining low import tariffs.

We establish two main results. The first result is a moderate role of constituent interests. An interquartile difference in the value of the the China shock decreases the probability of voting in favor of NTR for China by roughly 3-5 percentage points, while an analogous difference in ideology creates a 13-17 percentage point increase in the probability of supporting NTR for China. In our heterogeneity analysis, we show that constituent interests are more important for Democrats than for Republicans, and for politicians that were elected with small vote margins (a margin of responsiveness supported by other studies, see discussion in Mian et al. 2010; Ladewig 2010).

The second main result of our analysis is that politicians possessed a significant amount of knowledge about the future China shock. In some years we cannot reject that they perfectly forecasted the shock that would hit their district in the next five years. More precisely, for all years from 1990 to 2001 we cannot reject that politicians had, at least, enough information to

---

[3]There is a vast literature discussing pure economic models of voting where electoral constituents (Peltzman, 1984) or subconstitutents matter for roll call voting in Congress versus ideology of members of Congress (Kalt and Zupan, 1984, 1990; Levitt, 1996). For a recent review see (Mian et al., 2014).

forecast 55 percent of the variation in the China shock. Perhaps surprisingly, our findings imply that knowledge decreases over the 1990s, a result that is plausible given that China's comparative advantage shifted substantially during the late 1990s and early 2000s. Comparing legislators across parties, we find that Democrats were systematically more informed than Republicans, with the exception of the early 1990s, a period when they are both equally informed.[4] We also present several validation exercises for our approach, including placebo exercises and a comparison of the NTR voting for China to the NTR voting for the case of Vietnam, showing how the moment inequality estimation highlights similar patterns in terms of information sets and preferences for comparable votes. Our findings on the extent of the information sets of U.S. politicians (and their fairly accurate expectations) appear in line with the extent of information inferred from stock price responses around the China permanent NTR vote (e.g. Greenland et al., 2020).

Finally, we employ the estimated model to perform counterfactual exercises in which we give politicians perfect information about the upcoming shocks and calculate the change in voting behavior that would have resulted from the additional information. We find that overall support for China's NTR status would not have changed substantially in the presence of perfect information.

In essence, to the question in the title "Did U.S. politicians expect the China shock?" our answer is "Yes, but they did not give it substantial weight." Counterfactual simulations in section 6.2, where such weight is increased for all lawmakers in our sample for given baseline information, show that pro-China legislation would have been overturned.

One important premise to the question we are posing is the assumption that the electorate is generally attuned to trade policy positions of their representatives. It would otherwise be unclear why politicians would care about the reaction of voters. While it is implausible to assume that voters have a complete command of specific trade policy measures, a number of papers have documented the impact of the China shock and the recent trade war on electoral outcomes. Autor et al. (2020) find that districts more affected by the China shock saw an increase in Fox News viewership and elected more conservative Republicans and, to a lesser extent, more liberal

---

[4]Anecdotally, the Congressional Record reports statements by members of Congress about the expected labor market consequences of the China shock which turned out to be fairly accurate. For instance, during the permanent normal trade relations debate in 2000, Congress members David Bonior (D-Michigan) and Barbara Lee (D-California) shared predictions from the Economic Policy Institute, a nonprofit, for the state of California: "*In my State of California we estimate 87,294 jobs lost...*" over the next decade (https://www.congress.gov/congressional-record/volume-146/house-section/page/H3157). Similarly, but citing data from the International Trade Commission, Bill Pascrell (D-New Jersey) went on record stating: "*In New Jersey, we will lose 23,000 jobs. In the United States as a whole, we will suffer a net job loss of 872,000 jobs over the same 10 years. We are not creating jobs in America, we are creating jobs in China.*" (https://www.congress.gov/congressional-record/volume-146/house-section/page/H3514). Members of Congress acquired the information, among others, from labor unions (and especially so members of the Democratic Party). As Baldwin and Magee (2000) (p.83) state, around the debate over NAFTA and China NTR, "*most labor unions were convinced that the adverse employment effects would be much more widespread than economists had predicted*" and therefore the expectations of some of these members of Congress differend from part of the economic consensus at the time.

Democrats, thus inducing more polarization.[5] Another recent contribution by Che et al. (2020) finds that the 2010's reaction to the China shock was due to the anti-trade turn taken by the Republican party after the appearance of the Tea Party. Blanchard et al. (2019) also document a significant electoral impact of the trade war on the vote share of Republicans in 2018. Interestingly, the negative effect on GOP vote share coming from retaliatory tariffs imposed by U.S. trading partners is not mirrored by a positive effect due to the protection offered by import tariffs. In sum, these recent papers offer a clear justification for making constituent interests a major component of the decision to vote on an important trade policy measure, like maintaining and expanding China's Normal Trade Relations status.[6]

This paper contributes to the literature on Congressional voting, in particular on trade policy. An example is Baldwin and Magee (2000), which estimates the importance of constituent interests and campaign contributions from business and labor groups in three trade bill votes in the 1990s. Relative to Baldwin and Magee (2000), we sharpen the estimation of the role of constituent interests by employing a more precise measure of how constituents were impacted by the policy, but mostly by applying a new econometric methodology to the expectations of politicians. A more recent paper by Feigenbaum and Hall (2015) studies the impact of the China shock on congressional voting on trade bills in general and finds that congressmen in districts more negatively affected are less likely to vote in favor of trade promoting bills, as classified by the Cato Institute, a think-tank in Washington DC. Differently from Feigenbaum and Hall (2015)'s retrospective view, we take a prospective angle in modeling voting behavior, where law makers are deciding to vote based on the future electoral consequences of their decision. In the broader literature on the political economy of trade policy, Rodrik (1995) offers a more conceptual framework and depicts trade policy as emerging from demand (interest groups, grassroots, etc) and supply (government) factors. This paper's contribution sheds light on the individual behavior of legislators that constitutes a crucial element of the policy supply side.

Beyond the trade policy literature, this paper speaks to the established empirical literature focused on modeling voting in legislatures. The elements of this vast scholarship that are closer to our paper span political economy and political science (Poole and Rosenthal, 1984; Levitt, 1996; Heckman and Snyder, 1997; Poole and Rosenthal, 1997; Jenkins, 2000; Clinton et al., 2004; McCarty et al., 2006; Canen et al., 2020). Modeling prospective behavior of legislators does figure in this strand of research, as it is often postulated that a representative politician acts by "determining her roll-call vote choice based on which legislative options will maximize her future utility" (Ladewig, 2010). Stimson et al. (1995) argue that in congressional voting "elected politicians. . . sense the mood of the moment, assess its trends, and anticipate its consequences for

---

[5]Colantone and Stanig (2018) document a similar result for Western European countries.

[6]A recent paper by Fajgelbaum et al. (2020) points to the importance of constituent interests in the structure of tariffs in the Trade War the US started in 2018. U.S. import tariffs are such that marginal counties (those with a Republican vote share of around 50 percent) receive the highest level of protection.

future elections". However, expectations and information sets of lawmakers are rarely explicitly modeled as part of standard empirical approaches. The more complex exercise of assessing whether a politician may not be responding to prospective constituent conditions in their vote because of limited information or because of policy preferences appears unexplored. We contribute to this by showing how this moment inequality approach provides a useful stepping stone in estimating prospective behavior of lawmakers.

Less directly, our application offers a complementary view to the empirical literature focused on modeling the expectations of policy makers. This analysis has traditionally found important applications in Macroeconomics (Primiceri, 2006; Sargent et al., 2006), and in this sense, the paper connects to a broader set of questions than congressional voting alone. Our estimates also speak to the modeling of government preferences, a key area of political economy.[7]

2 Empirical model

This section presents a simple model of probabilistic voting for members of Congress. Indicate a congressional cycle with $t = 1, 2, ..., T$. Each period a single bill focused on a main policy issue is introduced – in this application maintaining Normal Trade Relations with China. As previously discussed, such bills were typically presented to the U.S. legislative branch and voted upon once per congressional cycle, so that t may equivalently indicate time and bill number.

Let us indicate with $x_t \in \mathbb{R}$ a policy position favorable to trade normalization, so that a Yes vote will indicate a vote for x_t. Consequently, one interprets a No vote as a vote against normal trade relations, $q_t \in \mathbb{R}$. This notation allows for the two positions to be affected by some nuance over time and neither position is assumed to be exactly constant in time. In the empirical analysis we will simply make sure that a Yes vote will be consistently labeled to the support for the alternative x_t.

Indicate by $i = 1, ..., N$ individual legislators, where N is large.[8] We will assume that individual i's preferences are described by a random utility framework. We also posit a spatial voting environment for the members of Congress.[9]

For simplicity of exposition (relaxed in the empirical application later), the deterministic component of the politician's utility is assumed to depend on: (i) the distance of the bill from his/her ideological position θ_i; (ii) an electoral motive, summarized by his/her expected future electoral support $V_{i,t+1}$ (for example, due to that expressive voters who are adversely impacted by the China

---

[7]For early applications, see Alesina (1988); Alesina and Tabellini (1990); Drazen and Masson (1994).

[8]What follows can be applied to each chamber independently at the cost of omitting interactions between the two chambers, such as resolutions and conferences. $N = 435$ for the House and $N = 100$ for the Senate.

[9]Spatial voting is a successful and informative modeling approach to the description of congressional behavior and it has found substantial support in the literature (Poole and Rosenthal 1997; Heckman and Snyder 1997; Clinton et al. 2004; McCarty et al. 2006; Bateman et al. 2017).

shock may reward or punish i based on his/her voting records.)

Concerning (i), political ideology $\theta_i \in \mathbb{R}$ is a unidimensional and fixed characteristic of i. The assumption of unidimensionality is appropriate in the time period under analysis (see McCarty et al., 2006). The assumption of constant policy preferences has been validated in the literature on congressional voting (Poole, 2007 for a discussion) and our results do not appear sensitive to replacing constant ideal points with time-varying ideal points using the estimates from Nokken and Poole (2004), which are available from the authors. We use the ideological positions θ_i from DW-Nominate first dimension scores (Poole and Rosenthal, 1997).[10] This follows a common approach in modeling congressional voting when the explicit estimation of such preference parameters is peripheral to the main empirical analysis like in this case (e.g., Mian et al. 2010, 2014).[11]

Concerning (ii), let us indicate by $S_{i,t}$ a 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 (the China shock as presented in Autor et al. 2016). Assume the potential electoral impact of the China shock in the district represented by politician i at time $t+1$ is defined by:

$$V_{i,t+1} = h_t(d_{i,t}, S_{i,t+1}) + e_{i,t+1},$$

where $d_{i,t}$ is the voting decision made by the politician, and

$$
\begin{aligned}
\mathbb{E}\left[e_{i,t+1}|d_{i,t}, S_{i,t+1}, \mathcal{I}_{i,t}\right] &= 0 \\
h_t(d_{i,t}, S_{i,t+1}) &= \gamma_t^0 + \gamma_t^2 S_{i,t+1} + \left(\gamma_t^1 - \gamma_t^2\right) S_{i,t+1} \times \mathbf{1}\left\{d_{i,t} = \text{vote for } x_t\right\}, \quad (1)
\end{aligned}
$$

where $\mathbf{1}\left\{\cdot\right\}$ is an indicator function and $\mathcal{I}_{i,t}$ is the information set of politician i at t. Note that the function $h_t(\cdot)$ introduces both a direct effect of the China shock on electoral support independently of i's vote and a component that depends on the interpretation by the voters of their representative i's decision. Appendix A.1 presents a full microfoundation of equation (1).[12]

---

[10]For further reference, see www.voteview.com

[11]The reader interested in the estimation of θ_i can find a detailed analysis in Canen et al. (2020) and references therein. Due to lack of sample overlap, the Canen et al. (2020) estimates cannot be used in our application.

[12]When considering the economic effects of trade with China, it may be natural to consider the role of exports, and not only imports. Recent work by Feenstra et al. (2019) shows that the exports-led increase in the demand for labor has almost matched the negative employment effects of the China shock. An important observation in this regard is that Feenstra et al. (2019) consider U.S. exports not only to China, but to all its trading partners. When considering only China as a destination market, exports and export growth were markedly smaller and did not have as large a positive effect on employment, as initially shown in Autor et al. (2013). Since NTR votes did not have obvious implications for the U.S. worldwide export prospects, we exclude exports from our main analysis of voting decisions, but include them in a robustness section in Appendix E.3.

The expected utility for a politician i of taking decision d_t, given information set $\mathcal{I}_{i,t}$ is:

$$U(\xi_{i,t}, d_{i,t}; \theta_i, \mathcal{I}_{i,t}) = u\left(\| d_{i,t} - \theta_i \|\right) + \tilde{\delta}\mathbb{E}\left[V_{i,t+1}|d_{i,t}, \mathcal{I}_{i,t}\right] + \begin{cases} \xi_{i,t,x} & if\ d_{i,t} = \text{vote for } x_t \\ \xi_{i,t,q} & if\ d_{i,t} = \text{vote for } q_t \end{cases},$$

where $u\left(\| \cdot \|\right)$ indicates an ideological loss that is function of the distance of the policy from the ideal point of i. The term $V_{i,t+1}$ indicates the future electoral outcome for the district represented by i. We assume a quadratic loss function $u(\cdot)$ and i.i.d. Gaussian term $\xi_{i,t,d} \sim N(0, \sigma_\xi^2)$. This implies the useful convolution $\xi_{i,t} = \xi_{i,t,q} - \xi_{i,t,x} \sim N(0, 2\sigma_\xi^2)$. A standard identification requirement in discrete choice problems with Gaussian shocks (e.g., in probit) further requires the normalization $2\sigma_\xi^2 = 1$, which we impose.

We define the variable $Y_{i,t}$ as an indicator function that is equal to 1 when legislator i decides to vote Yes on x_t and 0 when the legislator votes in favor of q_t :

$$
\begin{aligned}
Y_{i,t} &= \mathbf{1}\{U(\xi_{i,t}, x_t; \theta_i, \mathcal{I}_{i,t}) > U(\xi_{i,t}, q_t; \theta_i, \mathcal{I}_{i,t})\} \qquad\qquad (2) \\
&= \mathbf{1}\left\{ -\frac{1}{2}\left((x_t - \theta_i)^2 - (q_t - \theta_i)^2\right) + \tilde{\delta}\left(\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right]\right) \geq \xi_{it} \right\}.
\end{aligned}
$$

We can write the probability of $Y_{i,t} = 1$ as:

$$Pr(Y_{i,t} = 1|\mathcal{I}_{i,t}) = \Phi\left(\begin{array}{c} -\frac{1}{2}\left((x_t - \theta_i)^2 - (q_t - \theta_i)^2\right) \\ +\tilde{\delta}\left(\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right]\right) \end{array} \right), \qquad (3)$$

where Φ is the standard normal cumulative density function and $\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right]$ is the expected net loss (or net gain) of electoral support due to the China shock in the constituency represented by politician i in the future electoral cycle, given the information available to i at t. This implies that the probability of voting Yes depends on the expectations of the electoral consequences of voting Yes relative to voting No, which are unobserved by the econometrician. Intuitively, the higher is the relative expected electoral gain of voting Yes, the higher the likelihood of voting Yes.

It follows from (1) that:

$$\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right] = \left(\gamma_t^1 - \gamma_t^2\right)\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right].$$

Setting $\delta_t = \tilde{\delta}\left(\gamma_t^1 - \gamma_t^2\right)$ and simplifying the relative loss function $-\frac{1}{2}\left((x_t - \theta_i)^2 - (q_t - \theta_i)^2\right)$ as $a_t\theta_i + b_t$, where $a_t = x_t - q_t$ and $b_t = \frac{1}{2}(q_t^2 - x_t^2)$, we rewrite (2) and (3), respectively, as:

$$
\begin{aligned}
Y_{i,t} &= \mathbf{1}\{a_t\theta_i + b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right] \geq \xi_{i,t}\}, & (4) \\
Pr(Y_{i,t} = 1|\mathcal{I}_{i,t}) &= \Phi\left(a_t\theta_i + b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]\right). & (5)
\end{aligned}
$$

This structure of preferences separates (in a somewhat restrictive way) the role of a_t, interpretable as the politicians' weight on ideology, defined as an aggregate across all primary issues (e.g. taxes, gun control, government deficit, etc.) of concern to them and/or to local voters, and the role of δ_t, specifically indicating the differential alignment with constituents along the future China Shock dimension once the role of average ideology is accounted for. We discuss two distinct approaches to the estimation of these and the other parameters in the next section.

3   Expectations and information set of politicians: Estimation

A key contribution of this paper is the analysis of the information set available to politicians to forecast the labor market effects of the China shock at the time of a roll call vote. There are two fundamentally different approaches, which in turn hinge on the answer to the following question: is the politician's information set $\mathcal{I}_{i,t}$ known to the econometrician? When the answer is in the affirmative, then estimation can be performed by maximum likelihood or method of moments. When the econometrician knows only a subset of the information available to politicians, then one can adopt a moment inequality estimator. We discuss these two approaches in turn, but we first start with describing the three benchmark information sets that we will consider throughout the paper:

(i) Minimal Information: the politician knows his own ideological position θ_i, but the only information a politician has about the economic impact of the China shock is the current share of population employed in manufacturing in district i, $ShareMfg_{i,t}$;

(ii) Baseline Information: the politician has access to the minimal information set, plus the current period China shock $S_{i,t}$;

(iii) Perfect Foresight: the politician has perfect foresight of the labor market consequences of the China shock, so that $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right] = S_{i,t+1}$.

3.1   Politician's information set fully known to the econometrician: MLE

When the econometrician knows the content of the politician's information set, then the parameter vector $\omega_t = \{a_t, b_t, \delta_t\}$ can be estimated by maximum likelihood for each cycle. Based on expression

8

(5), the log-likelihood function takes the form:

$$
\ln \mathcal{L}\left(\omega_t | \{Y_{i,t}, \theta_i, \mathcal{I}_{i,t}\}_{i=1}^N\right) = \sum_{i=1}^N Y_{i,t} \ln\left[\Phi\left(a_t \theta_i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1} | \mathcal{I}_{i,t}\right]\right)\right]
$$
$$
+ (1 - Y_{i,t}) \ln\left[1 - \Phi\left(a_t \theta_i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1} | \mathcal{I}_{i,t}\right]\right)\right]. \tag{6}
$$

Maximizing (6) requires specifying the information set $\mathcal{I}_{i,t}$. In the case of perfect foresight, (6) is maximized after replacing $\mathbb{E}\left[S_{i,t+1} | \mathcal{I}_{i,t}\right]$ with $S_{i,t+1}$. In the case of minimal information set, the expectation of the China shock is derived as the predicted value of the following OLS regression: $S_{i,t+1} = \beta_0 + \beta_1 \theta_i + \beta_2 ShareMfg_{i,t} + \varepsilon_{i,t+1}$.[13] For the baseline information set, we can perform a similar two-step procedure, albeit with an OLS regression that contains a larger set of regressors, reflecting a richer knowledge by the politician. This methodology also imposes that politicians have rational expectations, i.e. a mean zero expectation error that is uncorrelated with the expectation.

The key assumption of the maximum likelihood approach is that we, as econometricians, are confident about what enters the politician's information set. When one misspecifies the politician's information set, the parameter estimates ω_t will be biased. The direction of the bias cannot be characterized in general, so for our case we resort to Monte Carlo simulations to illustrate the problem in Appendix B.

One specific instance lends itself to an intuitive explanation. When the econometrician incorrectly assumes that the politician has perfect foresight, the bias that arises is similar to the case of error in variables in a linear regression setting. The intuition is that $\mathbb{E}\left[S_{i,t+1} | \mathcal{I}_{i,t}\right]$ is measured with error when we replace it with $S_{i,t+1}$ and that error is, by assumption of rational expectations, uncorrelated with $\mathbb{E}\left[S_{i,t+1} | \mathcal{I}_{i,t}\right]$. Similarly to a linear regression setting, this will lead to an attenuation bias in the estimated coefficient δ_t. Assume that the true δ_t is negative and consider two representatives in districts A and B, who form their expectations based only on a minimal information set, which includes the manufacturing share in the region. Assume that district A and B have similar manufacturing shares, but different industrial composition. Hence, the two representatives predict a similar import shock, but in reality district A is much more severely affected than district B. We, the econometricians, assume that these representatives are instead very well informed about the imminent import increases. Because A and B expect a similar impact, they vote similarly on the bill. The econometrician, however, observing a similar voting behavior between politicians A and B, concludes that δ is smaller (in absolute value) and that the politicians place little weight on the import shock. The bias can be large (around 40%) under realistic data configurations, as shown in Appendix B.

---

[13]See Manski (1991) and Ahn and Manski (1993).

3.2 Politician's information set partially known to the econometrician: Moment inequality approach

In the previous subsection we have shown that the maximum likelihood approach relies on an accurate knowledge by the econometrician of the information set possessed by the politician. The alternative estimation method proposed by Dickstein and Morales (2018), based on moment inequalities, does not require full knowledge of $\mathcal{I}_{i,t}$, but rather of a subset of variables $Z_{i,t} \subseteq \mathcal{I}_{i,t}$. That is, the politician may know more than $Z_{i,t}$ in forming his/her forecast, but she knows at least the covariates in $Z_{i,t}$. Assuming only partial knowledge of $\mathcal{I}_{i,t}$ comes at the cost of less precise identification. We will not be able to point identify the elements of the parameter vector ω_t, but only to set identify them. Whether these sets are sufficiently tight to be informative will be carefully discussed in the results section.

The second important goal of our analysis is to ascertain the extent of the information set of legislators. This, in turn, involves a formal analysis of which subset of variables a politician considers at the time of his/her vote through an application of specification selection tests proposed by Bugni et al. (2015). Being able to reject that certain variables are used in the politician's forecast allows us to learn about the process of decision making of legislators: what they knew and considered relevant at the time of their vote. In this exercise the voting model and data are kept constant, but the subset of variables assumed part of the information set is varied.

In what follows, we allow politicians to have time varying information sets and we formally test whether certain groups of legislators have identical information sets or not. For instance, we assess whether members of higher levels of chamber seniority have broader information sets than lower seniority members, or whether members of opposing parties share the same information set. Questions of asymmetry of information sets across party lines are increasingly common in the political economy literature focused on polarization[14] and our application offers a formal approach to this problem for members of Congress.

A final question that the approach allows us to answer is whether, had politicians had a more complete information set, their votes for trade normalization with China would have been different. These counterfactuals are simulated within the same structure of expectations and information we just described.

Throughout, we maintain the assumption of rational expectations on the part of politicians, that is the expectational error $\varepsilon_{i,t+1} = S_{i,t+1} - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$ has mean zero, $\mathbb{E}\left[\varepsilon_{i,t+1}|\mathcal{I}_{i,t}\right] = 0$ and is uncorrelated with $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$. This means that politicians do not systematically skew their prediction or ignore elements of their information set which would systematically help in forecasting $S_{i,t+1}$. In Appendix A.2, we discuss different plausible data generating processes that support the rational expectation assumptions, given the formulation of the baseline China shock

---

[14]See Alesina et al. (2020).

measure.[15] We follow Dickstein and Morales (2018) in generating two sets of moment inequalities that identify the possible values that the parameters of interest can take: Odd-based moment inequalities and Revealed Preference moment inequalities.[16] In the following subsection we go through the main steps of the derivation of the inequalities to illustrate the basic intuition.

3.2.1 Odds-based moment inequalities

We use the definition in (4) to obtain:

$$\mathbf{1}\{a_t\theta_i + b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right] - \xi_{i,t} \geq 0\} - Y_{i,t} = 0. \tag{7}$$

This expression depends on the unobserved shock realization $\xi_{i,t}$ and $\mathcal{I}_{i,t}$. Therefore, we take the expectation of (7) conditional on $\mathcal{I}_{i,t}$ and manipulate the expression to obtain the following equality:

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

This equality still depends on the expectation $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$, which in turn depends on the true information set $\mathcal{I}_{i,t}$, an object that we do not observe. However, under the assumption that the expectational error $S_{i,t+1} - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$ has mean zero and from the property that $\frac{\Phi}{1-\Phi}$ is convex, one can replace $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$ with $S_{i,t+1} - \varepsilon_{i,t+1}$ and apply Jensen's inequality to derive the following inequality:

$$\mathbb{E}\left[(1 - Y_{i,t})\frac{\Phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)}{1 - \Phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)} - Y_{i,t}\middle|\mathcal{I}_{i,t}\right] \geq 0 \tag{8}$$

Consider now a subset of the information set $Z_{i,t} \subseteq \mathcal{I}_{i,t}$. By invoking the Law of Iterated Expectations, we may replace the unobserved information set $\mathcal{I}_{i,t}$ by $Z_{i,t}$, and obtain the following inequality from (8):

$$\mathbb{E}\left[m_l^{ob}\middle|Z_{i,t}\right] \geq 0 \tag{9}$$
$$m_l^{ob} = (1 - Y_{i,t})\frac{\Phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)}{1 - \Phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)} - Y_{i,t}$$

Notice that (9) is increasing in δ_t, so this condition identifies a lower bound for this parameter. We use the subscripts l, u to indicate the lower bound and upper bound inequalities.

Following a similar logic, one can derive a moment condition that further bounds the parameters of interest ω_t:

---

[15]We also assess the implications of different types of violations of the rational expectations assumption in Appendix F.

[16]Specifically, see their Appendix C for additional details.

$$\mathbb{E}\left[m_u^{ob}\middle|\,Z_{i,t}\right] \geq 0 \tag{10}$$

$$m_u^{ob} = Y_{i,t}\frac{1 - \Phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)}{\Phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)} - (1 - Y_{i,t})$$

Notice that m_u^{ob} is decreasing in δ_t and therefore moment inequality (10) identifies an upper bound for this parameter.

3.2.2 Revealed preference moment inequalities

The second set of moment inequalities derives from the revealed preference argument that a politician will vote Yes if and only if the benefit from doing so is positive, hence:

$$Y_{i,t}\left(a_t\theta_i + b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right] - \xi_{i,t}\right) \geq 0 \tag{11}$$

Because, again, $\xi_{i,t}$ is unobserved, we take the expectation of (11), conditional on $\mathcal{I}_{i,t}$ and obtain the following inequality:

$$\mathbb{E}\left[Y_{i,t}\left(a_t\theta_i + b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]\right) + \Gamma_{i,t}\middle|\,\mathcal{I}_{i,t}\right] \geq 0 \tag{12}$$

where $\Gamma_{i,t} = -\mathbb{E}\left[Y_{i,t}\xi_{i,t}|\mathcal{I}_{i,t}\right] = (1 - Y_{i,t})\frac{\phi(a_t\theta_i+b_t+\delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}])}{1-\Phi(a_t\theta_i+b_t+\delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}])}$, and ϕ is the standard normal probability density function. Once again the expression in inequality (12) contains the unobserved expectation $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$ and information set $\mathcal{I}_{i,t}$. Because $\frac{\phi}{1-\Phi}$ is convex, with the rational expectation assumption and $Z_{i,t} \subseteq \mathcal{I}_{i,t}$, we can apply the same logic as for inequality (9). The resulting inequality will be weaker than (12) and is given by:

$$\mathbb{E}\left[m_l^{rp}\middle|\,Z_{i,t}\right] \geq 0 \tag{13}$$

$$m_l^{rp} = Y_{i,t}\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right) + (1 - Y_{i,t})\frac{\phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)}{1 - \Phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)}$$

Starting from another revealed preference inequality:

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

we can obtain a second revealed-preference moment inequality in a similar manner:

$$\mathbb{E}\left[m_u^{rp}\middle|\,Z_{i,t}\right] \geq 0 \tag{14}$$

$$m_u^{rp} = -(1 - Y_{i,t})\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right) + Y_{i,t}\frac{\phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)}{\Phi\left(a_t\theta_i + b_t + \delta_t S_{i,t+1}\right)}.$$

3.2.3 Partial Identification

The moment inequalities defined by (9), (10), (13) and (14) are conditional on values of the vectors $Z_{i,t}$, which we allow to contain different variables, characterizing different possible information sets possessed by politicians. The following theorem indicates that the true parameter vector $\omega_t = \{a_t, b_t, \delta_t\}$ is contained in the set of parameters that are in compliance with the odds-based and revealed preference moment inequalities. Hence, the parameters of interest are partially identified:[17]

Theorem. *(Dickstein and Morales, 2018) At the true value of the parameter vector $\omega_t = \{a_t, b_t, \delta_t\}$ the following four moment inequalities are satisfied:*

$$
\begin{cases}
\mathbb{E}\left[m_u^{ob}\middle| Z_{i,t}\right] & \geq 0 \\
\mathbb{E}\left[m_l^{ob}\middle| Z_{i,t}\right] & \geq 0 \\
\mathbb{E}\left[m_u^{rp}\middle| Z_{i,t}\right] & \geq 0 \\
\mathbb{E}\left[m_l^{rp}\middle| Z_{i,t}\right] & \geq 0
\end{cases}
$$

where $Z_{i,t} \subseteq \mathcal{I}_{i,t}$ is the set of variables known by politician i at time t.

Conditional moments (9), (10), (13) and (14) 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 moment inequalities into unconditional moment inequalities, which can be directly employed in estimation. 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 again an issue to be reckoned with once we discuss our results.[18] The estimation implementation is detailed in Appendix C.

3.3 Further robustness of the methodology

A relevant issue pertinent to our application is whether politicians are uncertain about the impact of future import shock on future electoral support, and need to form expectations about it, as well as the China shock. This is an issue related to the uncertainty specific to the component $(\gamma_t^1 - \gamma_t^2)$ in the model, to which we need to explore sensitivity in the construction of our estimator. In

---

[17]See Appendix C in Dickstein and Morales (2018) for the proof of this result.

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

Appendix A.3 we clarify under which conditions we can allow this uncertainty and we reinterpret the coefficients estimates in light of this modification.

4   Institutional background and data

The background for the series of roll call votes that we employ in this paper is the extension of Normal Trade Relations to the People's Republic of China, after their suspension in 1951. NTR status was restored in 1980 under Title IV of the Trade Act of 1974 and was dependent on the presence of a bilateral trade agreement to be renewed every three years and on compliance with the Jackson-Vanik amendment on freedom of emigration, required for non-market economies.[19] China's NTR status would be renewed automatically every year upon the President recommendation unless Congress disapproved it by enacting a joint resolution. It is widely recognized that these resolutions were spurred by humanitarian and foreign policy considerations following the Tiananmen Square events of 1989 (Pregelj, 1998). Congress sought to provide incentives, through withholding of NTR status, to the Chinese government to address issues of human rights. In this effort, it clashed with the executive branch, a fact reflected in the several episodes in which the resolutions to disapprove NTR passed in the House, but died in the Senate or were overturned by a Presidential veto. In light of these considerations, it should be clear that we do not view the threat to local economic interests as the only driver of the legislators' roll call votes, and ideological considerations in the utility function of legislators account for this.

4.1   Roll call votes

The sample for our estimation includes individual legislator roll call votes for 12 House Joint Resolutions that took place every year from 1990 to 2001, as listed in Table D.1. Three of these Joint Resolutions to disapprove NTR extension were passed by the U.S. House of Representatives[20] in the years 1990, 1991 and 1992, but not voted on or struck down in the Senate.[21]

The website voteview.com provides the roll call votes, together with the ICPSR code for each legislator, the congressional district, the Party, and the first two dimensions of the DW-Nominate score, a multidimensional scaling application developed by Poole and Rosenthal (1997) and the first dimension of which is our proxy for θ_i.[22] Instead of "Yea" and "Nay", we indicate all votes

---

[19]See CRS Report for Congress (Pregelj, 2001) for further details.

[20]In what follows we exclusively focus on votes in the House of Representatives to exploit the pertinent commuting zone level variation in the local labor market effects of the China shock.

[21]Two House Resolutions, HR2212 and HR5318, in the 102nd Congress passed both in the House and the Senate and were vetoed by President G.H.W. Bush. There was no action on NTR bills of China in the Senate after 1992.

[22]voteview.com represents one of the most comprehensive and popular sources of measures of ideological positions in American politics and is a standard reference in the political economy literature on Congress. See the Introduction for references.

as Pro and Against China (NTR) for ease of interpretation (a "Yea" vote in favor of disapproving China's NTR is a vote against China).

Figure 1 shows that support for NTR is not purely along party lines and changes over time. Democrats are relatively more supportive of NTR in the middle of the sample period, while Republicans become increasingly supportive of NTR over time. There is switching of positions within individual legislators as well. Figures 2 and 3 show, by party, how many legislators switch position or maintain their vote relative to the previous year. On average, every year 15 percent (17 percent) of Republican (Democratic) legislators change their position relative to the previous year. In sum, there is sufficient heterogeneity in positions across parties and within legislators over time to justify that the changes in the electoral effects of the vote could play a role beyond constant preferences and ideology of legislators.

We also include in the data the voting outcomes of the bill HR 4444 in 2000 which would grant China permanent normal trade relations (PNTR) conditional on China's accession to the WTO. The results are stable regardless of the inclusion or exclusion of this bill.[23]

Note that the period 1990-1992 covers the final years of the George H.W. Bush administration.[24] The period 1993-1996 coincides with the first Bill Clinton administration and the the period 1997-2001 covers the second Clinton administration and the first year of the George W. Bush presidency. Given the important role played by the executive branch in the legislative evolution of China's NTR status, this subdivision of the sample period is explored in our analysis, by allowing for different parameters and expectations during each administration.

4.2 The China shock

The exposure to the China shock at the district level is generated from the import shocks in different local labor markets nested within the district. We start with constructing the import shocks at the level of commuting zones (CZs), which are clusters of adjoining counties characterized by strong commuting ties and have been conceptualized as local labor markets in the literature (Autor et al., 2013; Acemoglu et al., 2016). The shocks from different CZs are then aggregated to congressional districts (CDs) as in Autor et al. (2020).

---

[23]The shares of votes in favor of China by year is reported in Figure 4.

[24]NTR votes became not perfunctory only after the historical event of Tienanmen Square. From 1979 through 1989, presidents Carter, Reagan and Bush did not face any congressional opposition to the granting of Most Favored Nation privileges to China. (Johnson, 2005)

4.2.1 Exposure to import shock at the commuting zone level

Future supply shocks from China faced by the commuting zone j is constructed according to:

$$S_{j,t+1} = \sum_k \frac{L_{jk,t}}{L_{j,t}} \frac{\Delta M_{k,t+1}^{oth}}{Y_{k,t} + M_{k,t} - X_{k,t}}. \tag{15}$$

In this expression, $\Delta M_{k,t+1}^{oth}$ is the change in import of good k from China by eight other (non-U.S.) high-income countries over 5 years in the future.[25] It reflects the rising supply capacity of China due to its economic reforms, and is arguably exogenous to the U.S. product-demand shocks from the perspectives of the U.S. local economies (Autor et al., 2013; Acemoglu et al., 2016; Autor et al., 2020). The future import growth is then normalized by the contemporaneous absorption (U.S. industry output plus net imports, $Y_{k,t} + M_{k,t} - X_{k,t}$) at the industry level. $L_{jk,t}/L_{j,t}$ denotes the share of industry k in CZ j's total employment in period t.

The Bartik-style measure (15) summarizes the exposure of CZ j to China's future supply shocks from the standpoint of t. Having a perfect foresight of (15) not only requires the information on contemporaneous employment composition of the local labor market and domestic absorption of different industries, but also knowledge on supply shocks from China 5 years in the future. In our analysis, the future shocks correspond to the import supply growth over the period 1990-1995, 1991-1996, ..., 2001-2006, which overlaps with China's post-WTO-accession period when the U.S. witnessed the most intense increase in import competition from China.[26]

While the politicians may not have full knowledge of future import shocks as in (15), they may use the information of the past shocks to form expectations. For the years 1993-2001, we construct import shock in the past 5 years analogously as follows:

$$S_{j,t} = \sum_k \frac{L_{ik,t-5}}{L_{i,t-5}} \frac{\Delta M_{k,t}^{oth}}{Y_{k,t-5} + M_{k,t-5} - X_{k,t-5}}, \tag{16}$$

where $\Delta M_{k,t}^{oth}$ denotes the change in import of good k from China by eight other (non-U.S.) high-income countries over the previous 5 years.[27]

The baseline measures (15) and (16) follow the specification in Acemoglu et al. (2016) and

---

[25]The eight other high-income countries are Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland.

[26]For the baseline analysis, we construct future and past China shocks based on a 5-year window with the consideration that it is probably better able to reflect the underlying shift in China's import supply capacity while being consistent with the expected career horizon of the incumbents. Our results remain robust to alternative measures based on a 2-year window or a 10-year window (available upon request).

[27]Due to data constraints, for years 1990-1992, we use the 2-year-lagged variables to construct the past shocks. To be specific, for $t = 1990, 1991, 1992$

$$S_{i,t} = \sum_k \frac{L_{ik,t-2}}{L_{i,t-2}} \frac{\Delta M_{k,t}^{oth}}{Y_{k,t-2} + M_{k,t-2} - X_{k,t-2}},$$

Autor et al. (2020), and can be derived from workhorse trade models with a gravity structure. However, differently from the literature focusing on the impacts of contemporaneous trade shocks on local economies, our study aims at evaluating the extent to which politicians foresaw the future import shock from China in (15), and whether they acted on the relevant information when setting China-specific trade policies. The trade, employment and output data that are employed to construct the China shock measures are detailed in Appendix D.1.

4.2.2 Exposure to import shocks at the congressional district level

Following Autor et al. (2020), we map economic outcomes in CZs to CDs as follows. We start with the geographic relationships between counties and congressional districts provided by the Missouri Census Data Center (MCDC).[28] Counties are sometimes split across different CDs, and the MCDC concordance provides information on the distribution of the county population in each CD. We then ascribe to each county-by-congressional district cell the CZ-level import shock that corresponds to the county, and weight each cell by its share of population in the district. Lastly, we aggregate the weighted shocks across cells to the CD level. By construction, if a district spans multiple CZs, its exposure to China's rising import competition is the population-share-weighted average of the import shocks in these CZs. Since congressional districts, by construction, have similar population size, they have roughly the same weight in our analysis. In the empirical analysis, we denote the future and past import shocks at the CD level by $S_{i,t+1}$ and $S_{i,t}$, respectively.

During the sample period 1990-2001, the boundaries of county-by-congressional district cells experienced a major change in 1993, but remained stable afterwards. Therefore, for 1990-1992, we map the CZ-level import shocks to congressional districts as defined for the 102nd Congress. For 1993-2001, the mapping is based on the configuration of the 103rd Congress. This treatment does not affect the consistency of our baseline analysis because, as is discussed below, we conduct the estimation by periods based on presidential administrations, and none of the subsamples spans over 1992-1993.

Figure 4 shows the cross-district averages of past and future import shocks. The import shocks are always positive throughout the sample period, but the future shocks move less in tandem with the past shocks in the later years. For the moment inequality estimation discussed in the following section, we detrend the import shocks. The corresponding distributions reported in Table D.2 reveal a substantial heterogeneity in exposures across districts. For the periods 1997-2001,

---

where ΔM_{kt}^{oth} denotes the change in import of good k from China by eight other (non-US) high-income countries over the past 2 years. As is discussed in Appendix D.1, we use the data from County Business Patterns (CBP) to construct employment shares. For 1990-1992, calculating $L_{ik,t-5}/L_{i,t-5}$ requires the 1985-1987 CBP data with industries classified based on 1977 SIC codes. For the purpose of analysis, the data needs to be mapped to the 1987 SIC codes. However, the crosswalk from 1977 SIC to 1987 SIC involves many splits of industries. As a result, the concordance leads to a structural break in the employment measures for some localities over 1987-1988. For the concern of systematic measurement errors, we don't use the CBP data prior to 1988 for the main analysis.

[28]http://mcdc.missouri.edu/applications/geocorr1990.html

1993-1996, and 1990-1992, the interquartile ranges of future shock are 0.154, 0.078, and 0.143, respectively.

5   Results from congressional voting on NTR with China

5.1   Estimation results

This section reports our main results. To benchmark our approach to more standard methods using incorrect proxies for the information set of politicians, we start by reviewing estimates of a voting model using a maximum likelihood approach.

Table 1 shows that results substantially vary depending on the informational assumptions made. An important parameter of interest is the weight placed by politicians on their affected subconstituencies δ_t. A reasonable prior for this parameter would be $\delta_t < 0$, i.e., a negative utility weight placed on electoral groups adversely affected by the China shock in the politician's congressional district (ceteris paribus, the politician wishes to minimize these adverse effects). While occasionally the parameter estimates for δ are negative and sometimes of magnitude similar to those obtained using the moment inequality approach, often the coefficients are economically insignificant and not statistically different from zero. For example, for the period 1997-2001 the estimate for δ is 0.018, which is small in absolute value relative to consistent estimates obtained with the moment inequality method. The risk of attenuation from misspecification of the information set of politicians is therefore evident in the MLE case. To further consolidate the intuition we also offer Monte Carlo evidence of the problem in Appendix B.

We now proceed to the results of estimating of equation (5) using the three information sets $Z_{i,t}$: Minimal; Baseline; and Perfect Foresight. The 95% confidence sets that we report are built through a grid search implementing the Generalized Moment Selection (GMS) method in Andrews and Soares (2010) as detailed in Dickstein and Morales (2018). For each value of ω_t we build a modified method of moments (MMM) statistic, which tends to be large when the moment inequalities are *not* satisfied at that value of the parameters. This is formally tested by constructing the asymptotic distribution of the MMM statistic and rejecting ω_t when the MMM statistic is above the critical value corresponding to the 95th percentile of that distribution. Incidentally, empty confidence sets instead have to be interpreted as highly significant rejection (p-value < 0.05) of the corresponding information set. The steps to construct the 95% confidence sets are detailed in Appendix C.2.

Table 3 reports estimation results splitting the NTR votes by presidential administration and for the full sample 1990-2001. The first period 1990-1992 covers the George H.W. Bush administration, the second period 1993-1996 coincides with the first Clinton administration, and the third period 1997-2001 covers the second Clinton administration.[29] Focusing on sub-periods allows for

---

[29]As a robustness check, we drop the year 2001, which overlaps with George W. Bush's first year in office during

heterogeneity in information and flexibility of the parameters with respect to the behavior/agenda setting of executive branch – a more robust approach that we prefer. We will mostly discuss results for the importance of constituency interests δ, but we will gauge its magnitude in relation to the role of ideology parameter a.

Consider the pooled 1990-2001 results first. Panel A reports estimation results under the assumption that the politicians' information set includes at least the manufacturing share in their district and their own ideology. Under this assumption we report a confidence set $[-1.163, -0.150]$, so that all values in the confidence set are negative, as we would expect if politicians are more likely to vote against China's NTR status when they expect their constituency to be exposed to a larger shock. The confidence set is quite similar under the assumption of baseline information, which includes also the past shock $S_{i,t}$, $[-1.438, -0.500]$.

The magnitude of the parameter δ can be illustrated as follows. Considering two districts whose value of the China shock are at the 25th and 75th percentile. The probability of voting in favor of China's NTR status decreases by between 0.022 and 0.059 when the value of the expected China shock goes from its 25th percentile to its 75th percentile.[30] This is a moderate effect, especially considering the role of ideology. If we perform the analogous exercise, we find that an interquartile range shift in ideology (from relatively liberal to relatively conservative) calculated at the mean expected China shock value produces an increase in the probability of voting in favor of China of between 0.134 and 0.175 in 1990-2001. In sum, these comparisons allow us to conclude that the effect of ideology is much larger than the effect of constituent interests. These results line up with common findings in the congressional voting literature for most bills (as early as Kalt and Zupan (1984) and for a recent discussion see Poole and Rosenthal (2017)).

We will discuss differences in the parameters estimates between subsets of legislators after introducing specification tests that allow us to gauge the information possessed by politicians.

5.2   Testing for different information sets

A fundamental first step in the analysis of politicians' decisions is to assess the exact extent of their information sets at the moment of the vote. Intuitively, when $Z_{i,t} \nsubseteq \mathcal{I}_{i,t}$, moment inequalities presented in subsection 3.2 are misspecified, and there could be no values of ω_t that yields the data moments consistent with the misspecified inequality conditions.[31] The estimated confidence set hence could be empty. A formal statistical approach to this form of specification test is

which the final vote took place, but was related to the previous administration's efforts (Pregelj, 2001). The results remain similar.

[30]In particular, we calculate these percentage points as the minimum and maximum of $\Phi\left(b + \delta\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}^b\right]^{75th}\right) - \Phi\left(b + \delta\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}^b\right]^{25th}\right)$ evaluated at the mean $\theta_i = 0$ and $\omega_t \in \Omega_t^{95\%}$, where $\Omega_t^{95\%}$ is the 95 percent confidence set of the underlying parameters.

[31]To see this, the derivation step that invokes the Law of Iterative Expectation is invalid. Hence, for some values of $Z_{i,t}$, the moment inequalities are unsatisfied for all values of ω_t.

19

presented in Bugni et al. (2015), which allows us to reject the null hypothesis that there exists a value of ω_t within the parameter space that can rationalize the set of odds-based and revealed preference-based moment inequalities.

Specifically, consider two alternative information sets $Z_{i,t}^1$ and $Z_{i,t}^2$. Suppose there is no value of the parameter vector for which the set of moment inequalities hold given $Z_{i,t}^1$, yet there are values of ω_t within the parameter space that can rationalize the set of moment inequalities given $Z_{i,t}^2$. Then one may infer that $Z_{i,t}^1 \nsubseteq \mathcal{I}_{i,t}$ and cannot reject $Z_{i,t}^2 \subseteq \mathcal{I}_{i,t}$. The rejection of the null in this framework may also indicate misspecification of the original model of decision, so simple rejection of $Z_{i,t}^1$ cannot exclude misspecification per se. What is crucial in this application is that the failure to reject $Z_{i,t}^2$ eliminates this second interpretation of the test. Misspecification of the original model would affect the analysis under both $Z_{i,t}^1$ and $Z_{i,t}^2$, as the decision model is unchanged and only the information set varies across tests, and model misspecification would imply rejection in both instances.

Appendix C.3 reports the full details for the construction of the BP, RC, and RS specification test statistics following Bugni et al. (2015) and the corresponding p-values. Generally speaking, the test BP is less powerful than RC and RS, and rejection of the null hypothesis in any of these tests indicates a rejection of the hypothesis that specific information belongs to the information set of the members of Congress.

We report results for all three tests in Table 3. Columns 4, 5, and 6 report the p-values for BP, RC and RS respectively. For the full sample and for two of the three presidential administration periods we reject that politicians have perfect foresight, with p-values of 0.01-0.015 depending on the test. For all periods we cannot reject that the politicians had at least the baseline information set. While it is not obvious whether the baseline information set represents the actual amount of knowledge, we believe Table 2 is informative. It shows that the manufacturing employment share and past import shock have significant explanatory power in predicting future shocks. The R-squared is generally around 55 percent and is lower in the later period, suggesting that the variance of the expectational errors have increased over time and close to the end of the sample.

Our results point to information used by politicians worsening over time and their capacity of forecasting the China shock in the following five years deteriorating. Why would legislators be better informed in the earlier part of the period? We hypothesize that this is due to the more predictable nature of the shocks in the early 1990s, when China was specialized in relatively less complex and more labor intensive products. In Figure 5 we report the autocorrelation of the China shock at the industry level and it is clear that in the earlier years the shock was more predictable from year to year as the autocorrelation is above 0.3. The decline in the autocorrelation parameter begins its negative adjustment in 1994, which matches (qualitatively) the timing observed for 1993-1996 and 1997-2001 in terms of heightened difficulty of predicting the China shock. This is compatible with the evolution of China's comparative advantage from low value-added to more

complex products over time.

Our evidence on the extent of the information set of U.S. politicians and our assessment of their (fairly accurate) predictions on the industry consequences of the China shock is in line with evidence concerning expectations of other types of actors. Greenland et al. (2020) find that equity valuations around the key PNTR vote in Congress correlated with U.S. firms' exposure to trade liberalization with China, supporting the view that stock market participants were systematically pricing firms and sectors exposure to the shock.[32] It is therefore not completely implausible that U.S. legislators had information sets comparable with those of financial agents.

5.3   Heterogeneity across groups of politicians

5.3.1   Party

The estimates and tests presented so far were performed in the universe of the members of Congress during the 1990-2001 period under an assumption of common information sets. It is plausible to hypothesize that politicians from different parties, and with different electoral prospects, might have had varying degrees of knowledge and different expectations.

In this section we analyze three dimensions of heterogeneity: political party, tenure in office (i.e., experience), and margin of victory in the most recent election. For all tests we report the results for the baseline information set. To anticipate our findings, the picture that will emerge from this heterogeneity analysis is that Democrats at the time of NTR votes were both more informed and more sensitive to constituent interests than Republicans, and that legislators in tighter electoral races placed a heavier weight on the China shock.

In Table 4 we find that Democrats are more informed than Republicans, as we cannot reject that Democrats had at least the baseline information set in all time periods, but we can reject at standard levels of significance that Republicans know the baseline information in all periods, but 1990-1992. We believe part of the rationale for this finding comes from the reliance of Democrats and some Northern Republicans on information from sources close to the labor movement.[33] When we obtain non-empty parameter estimates, we find that Democrats display higher (in absolute value) sensitivity to the China shock. Republicans' confidence set for δ straddles zero, while the entire confidence set for the Democratic members of the House is comprised of negative values.

---

[32]For a discussion on the degree of predictability of trade policy change effects based on stock reactions of both domestic and foreign firms see also Breinlich (2014).

[33]For example, the the Economic Policy Institute predicted in 2000 *"the elimination of 872,091 jobs during the next decade"*, due to the trade deficit with China induced by the permanent normal trade relations (PNTR). This is remarkably close to estimates subsequently produced by economists. Acemoglu et al. (2016) find that U.S. trade with China eliminated 985,000 American manufacturing jobs between 1999 and 2011. Autor et al. (2013) and Pierce and Schott (2016) estimate that the "China shock" reduced on net U.S. manufacturing employment by 1.5 million jobs between the year 1990 and 2007. Acemoglu et al. (2016) estimate close to a 1.98 million total jobs lost. See footnote 4 for some excerpts from the Congressional Record citing these sources.

In Table E.1, we estimate the model separately for Democrats and Republicans based on the specification with minimal information. In this case, we obtain non-empty confidence sets for both parties for each sub-sample period. For Republicans, the 95% CS of δ always contains zero. While the estimated confidence sets for Democrats become wider, it appears that the weight on constituent interests is smaller for Republicans compared to Democrats. These findings are in line with Democratic legislators historical alignment with workers' interests and placing more weight on the affected subconstituencies (Poole and Rosenthal, 1997).[34]

To further establish the relevance of this dimension of heterogeneity, we also considered behavior of politicians beyond voting, particularly congressional speech (number of speeches related to the "China and trade" issue or the "China and labor" issue delivered). We can see that the non-voting data aligns with preference and information patterns that we reported.

We only focus on two corroborating findings here and present a full analysis in Appendix E.1. First, as the China shock over 2001-2006 is realized, representatives from districts in the top tercile of the exposure to import shock from China raise the related trade and labor issues more often in their speeches, but such response is stronger for Democrats than Republicans. Second, consistently with their higher information and preference weights, Democrats start taking actions earlier. Specifically, for Democrats the congressional speech on China starts surging during the 106th Congress, 1999-2000, while for Republicans, the effect picks up in the 108th, 2001-2002.

5.3.2 Vote margins and tenure

In Table 5 we also explore whether politicians with above-median victory margins in the previous election display differences in the sensitivity to constituent interests. The literature has found higher sensitivity for legislators in tighter races (Mian et al., 2010). We indeed find that, across different periods, legislators in tighter races display confidence sets for δ that are entirely composed of negative values, whereas confidence sets for politicians in safe races often cover both positive and negative values, consistent with lower preference weights. It does not appear to be the case, however, that politicians in tight races were differentially informed relative to politicians elected by larger margins. While it is not an objective of this section to identify whether the heightened sensitivity to the China shock was due to state dependent preferences of politicians, changing with electoral conditions, or due to selection (although this should also reflect in different information sets), the analysis does display a potential for our approach to pick up differential elements of the behavior and knowledge of sub-groups of politicians.

Finally, we explore the role of experience. In Table 6 we divide the sample in two according to whether House members tenure is above or below the median. One may imagine that politicians that are very experienced have better access to various sources of information. We do not find

---

[34]In Appendix F.2 we further discuss through Monte Carlo simulations the role of misspecifications of information sets with respect to heterogeneity across levels of ideology θ.

support for this conjecture, as confidence sets for δ of junior legislators appear not systematically different from those of senior ones.

5.4   Role of special interests

Special interests' contributions[35] are often listed within the set of potential drivers of congressional voting, but not without substantial uncertainty about the economic magnitude of their effect. There is evidence of a prominent role of special interest giving in certain votes (e.g. the Emergency Economic Stabilization Act of 2008, Mian et al., 2010), but no consensus in the political economy literature on its role for the bulk of all congressional activity (Stratmann, 2005). In the case of China's NTR, we find no evidence that special interests, both in terms of campaign contributions from business organizations (corporations and business associations) or from labor unions, influence the estimated effects of ideology and constituent interests on congressional votes in our main specification. We base this assessment on three main sets of empirical evidence which we report below.

First, as an extra dimension of heterogeneity, we separate politicians into two groups, depending on whether the campaign contributions from business interests are above or below the median in the sample. The degree of heterogeneity found in Table 7 is minimal, with marginally more negative estimates of constituent weights for politicians with contributions above the median in later periods. This may suggest that money in politics may target politicians with some type of characteristics, but ultimately the confidence sets do not point to substantial differences.

Second, we augment our specification with campaign contributions. It has to be noted that adding elements to the vector of parameters within the moment inequality approach is extremely costly due to the grid search process necessary for inference and hypothesis testing. Further, contributions could be endogenous to NTR votes, and hence it may be difficult to interpret the additional parameters. Table E.2 in Appendix E reports the results. Due to the computational burden, we consider the following specifications, each with four parameters to estimate and information sets to assess: (a) Baseline specification + campaign contributions from business organizations (Panel A); (b) Baseline specification + campaign contributions from labor unions (Panel B). Our main results appear robust to these additional controls.

Third, we explore the congressional committees closer to the policy decision and likely the most important targets for special interests. In particular, we exclude from the analysis politicians working in influential committees (including the Commerce and Ways and Means committees),

---

[35]The data on campaign donations employed in the analysis of special interests are obtained from opensecrets.org, based on official Political Action Committee disclosure forms from the U.S. Federal Election Commission. opensecrets.org is a website run by the Center for Responsive Politics, one of the main nonpartisan organizations in Washington DC, dedicated to electoral transparency and to the collection of information related to campaign spending and lobbying disclosures.

who may be more influenced by money in politics. As is reported in Table E.3, our results remain robust to this approach.

5.5 Validation: NTR votes for Vietnam and votes on other foreign policies

As validation of our model and methodology, we briefly compare the results for the analysis of the NTR votes for China to a set of similar, but distinct NTR votes for the case of Vietnam. The goal here is to establish comparability in the responses of politicians across sets of votes. We view this as a way of supporting external validity of our findings. The analysis covers the votes on Vietnam's Jackson-Vanik Waiver (necessary to extend Vietnam's NTR status) that took place over the period 1998 to 2002. Based on a report by the Congressional Research Services, disapproval resolutions were not introduced in 2003, 2004, or 2005 (i.e. there is no voting data over 2003-05). In 2006, the House passed legislation to grant Vietnam permanent normal trade relations (PNTR) status as part of a more comprehensive trade bill and Vietnam accessed the WTO in 2007.[36]

In Table 8 the coefficient δ appears larger in magnitude than for the China shock case. Specifically, the baseline confidence set is $[-74.075, -6.700]$. However, this is because, as expected, the magnitude of the import shock from Vietnam is several orders smaller than that from China. The standard deviation of future import shock from Vietnam is 0.006, while that from China in the same period is 0.129. Adjusted for this scaling, the economic significance of constituency interests appears small for the Vietnam case as well, consistent with our findings of a low constituent weight estimated with the China NTR votes.

Concerning information sets, Table 8 confirms that members of Congress were informed about the impact of Vietnamese imports to some extent, as they were for China.[37] We cannot reject at standard significance levels the baseline information set, however, we reject that politicians have perfect foresight. Overall, across the sets of NTR votes for Vietnam and China, we do not find salient differences both in terms of economic magnitude of constituent weights and information sets.

As a final placebo check, we investigate the impact of the China shock on congressional voting on other foreign policies. In particular, we consider the annual Foreign Operations appropriations bill, which is the primary legislative vehicle through which Congress reviews the U.S. foreign aid budget. The appropriations bills introduced over 1990-2001 cover a wide range of foreign aid

---

[36]See Appendix D.2 for additional details on the data. Due to the congressional redistricting in 2002, for the analysis in this section, we only include the bills over 1998-2001. The past and future import supply shocks from Vietnam are constructed analogously to the China shock in section 4.2.

[37]The manufacturing employment share and past import shock also have significant explanatory power in predicting future shocks for the case of Vietnam. The R-squared is generally around 53 percent.

programs across different geographic regions. Hence – we conjecture – the voting decisions are unlikely to be driven by constituents' considerations related to the China shock and this should be picked up by our methodology. This is the case. In Table E.4, we find that the confidence sets of δ always contain zero across different periods, consistent with the conjecture that the China shock should have little influence on policies unrelated to issues on trade or local markets.

5.6   Further robustness

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. There are two main reasons for this modeling choice. First, it is empirically relevant, as politicians routinely campaign on their individual vote choices, and attack each other based on their respective individual voting records, rather than on the outcome of specific roll calls. Second, it is also a realistic assumption for decision making in the U.S. House where the set of agents has large cardinality. In our context, 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 evaluate the robustness of the baseline findings to the alternative modeling choice. Appendix E.2 proposes a simple pivotal voting model and re-formulate the measure of trade shock accordingly. In particular, the alternative measure incorporates the "NTR gap" as defined by Pierce and Schott (2016) and takes into account that votes to revoke China's NTR status would be more effective at reducing imports from China for goods where the difference between the Most Favored Nation tariff and the "Column 2" tariff is larger. Our baseline findings remain robust in this environment.

Appendix E.3 introduces export shocks and finds 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 our main findings.

Appendix F.1 discusses through Monte Carlo simulations the potential biases in our analysis due to violations of the rational expectation assumption. In particular, we consider the following scenarios: (i) constant under- or over-prediction of the China shock; (ii) expectational errors correlated with $S_{i,t}$; and (iii) expectational errors correlated with θ_i. In all these particular settings, the estimation based on moment inequalities and the specification tests appear to be robust to moderate violations of rational expectations. Exceptions typically emerge in cases where the variation in irrational expectational errors is larger than the variation in unpredictable components of the China shock (i.e., the component of $S_{i,t+1}$ that is unexplained by elements in $\mathcal{I}_{i,t}$).

6   Counterfactuals

In this section we employ the estimated parameters of the model to perform two counterfactuals that answer the following questions: (i) Would giving full information to legislators have changed the results of the NTR roll call votes? (ii) Would the results of the NTR votes have changed if legislators had placed a larger weight on the labor market consequences of the China shock?

6.1   Legislators receiving information about the China shock

After establishing that, for most of the time period in our sample, politicians had less than perfect knowledge of the labor market consequences of China's export expansion, a natural question we can ask is whether providing more accurate information to politicians would have changed their vote, and the overall passage of certain bills. The answer, for the case of China NTR votes, is not substantially.

Denote by $\mathcal{N}_t$ the set of politicians in period t, and $\Pi^+(\omega_t, \mathcal{I}_{i,t}, \mathcal{N}_t)$ the share of votes in favor of China. We simulate $\Pi^+(\omega_t, \mathcal{I}_{i,t}, \mathcal{N}_t)$ under the case of baseline information $\mathcal{I}_{i,t}^b$ and the case of perfect foresight $\mathcal{I}_{i,t}^p$. The change in share of votes (in percentage point) in favor of China when we provide politicians with full information is then given by:[38]

$$\left[\min_{\omega_t \in \Omega_t^{95\%}} \left\{ \Pi^+(\omega_t, \mathcal{I}_{i,t}^p, \mathcal{N}_t) - \Pi^+(\omega_t, \mathcal{I}_{i,t}^b, \mathcal{N}_t) \right\} \times 100, \max_{\omega_t \in \Omega_t^{95\%}} \left\{ \Pi^+(\omega_t, \mathcal{I}_{i,t}^p, \mathcal{N}_t) - \Pi^+(\omega_t, \mathcal{I}_{i,t}^b, \mathcal{N}_t) \right\} \times 100 \right].$$

Before we delve into the specific results, it is worth pointing out that the effect of information provision on NTR votes is ambiguous, and depends on (i) the underlying distribution of the expectational errors $G(\varepsilon_{i,t+1})$, (ii) the weight on constituent interests that is governed by δ_t, and (iii) the policy position relative to individual ideology $a_t \theta_i + b_t$. Regarding (i), our assumption of rational expectations dictates that $\varepsilon_{i,t+1}$ has a mean zero, conditional on $\mathcal{I}_{i,t}^b$. Moreover, as is shown in Appendix Figure E.2, the expectation errors are more or less symmetrically distributed in different sample periods. In Figure 6, we illustrate the roles of (ii) and (iii) given a symmetric distribution of $\varepsilon_{i,t+1}$. Point A represents the probability of casting a pro-China vote for legislators who are endowed with θ_i and have an expectation $\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b]$. Note that in this scenario, the general policy position is in favor of China, and point A is located in the concave segment of $\Phi(\cdot)$. (This scenario is likely the case during 1993-1996 when the vote share in favor of China was above 70 percent on average.) When these politicians are supplied with complete information, they face import shocks $S_{i,t+1}$ that are dispersed around $\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b]$ in a mean-preserving way. As is represented by point B, the share of pro-China vote will be lower in the counterfactual due

---

[38]As the construction of the counterfactual is technical and not commonplace in the literature, we provide its details in Appendix C.4.

to the concavity of $\Phi(\cdot)$ at this segment. The magnitude of the counterfactual change hinges on the dispersion of $\delta_t \varepsilon_{i,t+1}$, which in turn depends on δ_t and the variation of $\varepsilon_{i,t+1}$. In section 5, we show that the effect of constituent interests reflected by the estimate of δ is moderate, and that the covariates in the baseline information set have significant explanatory power in predicting future shock especially in the earlier period (i.e., the variation of expectational errors is also moderate). These findings indicate that the effect of improving information could be small. Points C and D in Figure 6 depict the case in which general policy position is against China, and point C is located in the convex segment of $\Phi(\cdot)$. (This is related to the scenario in 1990-1992 when the vote share in favor of China was around 40 percent on average.) Under this case, had the politicians been provided with complete information on $S_{i,t+1}$, the pro-China vote share would have increased. The magnitude of the counterfactual change again depends on δ_t and the variance of $\varepsilon_{i,t+1}$.

We also simulate the share of politicians in the following categories: (i) vote for China in the baseline, but switch vote in the counterfactual, (ii) vote in favor of China in both the baseline and the counterfactual, (iii) vote against China in the baseline and switch in the counterfactual, and (iv) vote against China in both scenarios.

Table 9 shows the results of this counterfactual exercise. There are only a few small changes in the voting patterns across politicians who now possess full information, a result of the fact that they already possess a substantial amount of information to begin with and that the weight on constituent interests is moderate. Moreover, the vote share in favor of China's NTR status appears unaffected.[39] In summary, the evidence shows that lack of information on the part of members of Congress was hardly a driver of the particular outcome of the NTR votes.

6.2   Counterfactual: Heightened constituent interests

Another counterfactual that we explore here involves increasing the legislator's weight placed on his or her local constituents. Specifically, we apply to all politicians the lower bounds of δ_t corresponding to the confidence sets of two groups of politicians (i) Democrats and (ii) politicians with below median victory margins in the previous election, and then simulate the NTR votes. As is demonstrated in section 5.3, both these groups place a higher weight on the subconstituencies affected by the China Shock. By setting δ_t at the highest possible (absolute) values, we assess whether this margin of preferences may have played an important role in driving legislative outcomes for given expectations of the legislators.

Table 10 reports the confidence sets of the simulated vote shares. For all the simulations, we assume that politicians have access to the baseline information set and make voting decisions based on $\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b]$. In Panel A, we simulate the shares of pro-China votes based on the baseline

---

[39]Also note that in aggregate the counterfactual change in pro-China vote share in 1993-1996 is negative, while that in 1990-1992 is positive, which is consistent with the scenarios depicted in Figure 6.

estimates reported in Table 3. The predicted vote shares in favor of China align with the actual shares.

In Panel B, we apply the largest possible weight that Democrats place on the expected China shock to all politicians. The counterfactual weight in 1997-2001 (respectively, 1993-1996 and 1990-1992) is nearly three times (respectively, two times and three times) larger than the lower bound of the baseline estimates of δ_t. The simulated vote share is in the range of $[17.485, 33.071]$ in 1997-2001, indicating that bills in favor of NTR with China would have not passed in this counterfactual scenario during this period. The only caveat is that the confidence set of the counterfactual vote share still straddles 50 percent for the period 1993-1996.

Panel C applies to all legislators the largest possible weight placed by the politicians facing high electoral pressure (i.e., winning margins below median). In this counterfactual, the passage of the bills in favor of NTR with China would have been overturned in periods 1993-1996 and 1997-2001.

7   Conclusion

China's permanent NTR status in the U.S. and its accession to the WTO represent one of the most salient critical junctures in international trade (and certainly in trade policy decisions) of the last fifty years. This paper investigates whether U.S. politicians had imperfect information about the extent of China shock's repercussions in their home district at the time when they repeatedly voted on China's Normal Trade Relations status between 1990 and 2001.

To isolate the role of preferences versus information of members of Congress, we present a voting model and an application of a method of moment inequality approach designed to estimate expectations in decision making and to formally test for the content of information sets of legislators. We find that U.S. legislators had imperfect, but fairly accurate expectations, yet placed a relatively low weight on the constituencies that ended up being adversely affected by the China shock.

The approach discussed in this paper may be a general and informative avenue for the study of policy decisions made by politicians and to understand the role of their expectations. Being able to resolve between preference and information margins is a valuable step forward for political economy and political science scholars interested in questions of both political accountability and learning in policy making. Future research should implement and extend our application outside the consequential questions of trade policy addressed in this paper. These could include quantifying the role of congressmembers' expectations for housing market reform ahead of the 2008-09 financial crisis, and labor, healthcare, and fiscal policy in the 2010's.

References

Acemoglu, D., Autor, D., Dorn, D., Hanson, G.H., Price, B., 2016. Import competition and the great us employment sag of the 2000s. Journal of Labor Economics 34, S141–S198.

Ahn, H., Manski, C.F., 1993. Distribution theory for the analysis of binary choice under uncertainty with nonparametric estimation of expectations. Journal of Econometrics 56, 291–321.

Alesina, A., 1988. Macroeconomics and politics. NBER macroeconomics annual 3, 13–52.

Alesina, A., Miano, A., Stantcheva, S., 2020. The polarization of reality, in: AEA Papers and Proceedings, pp. 324–28.

Alesina, A., Tabellini, G., 1990. A positive theory of fiscal deficits and government debt. The Review of Economic Studies 57, 403–414.

Andrews, D.W., Shi, X., 2013. Inference based on conditional moment inequalities. Econometrica 81, 609–666.

Andrews, D.W., Soares, G., 2010. Inference for parameters defined by moment inequalities using generalized moment selection. Econometrica 78, 119–157.

Ansolabehere, S., Snyder Jr, J.M., Stewart III, C., 2001. Candidate positioning in us house elections. American Journal of Political Science , 136–159.

Autor, D., Dorn, D., Hanson, G., 2019. When work disappears: Manufacturing decline and the falling marriage market value of young men. American Economic Review: Insights 1, 161–78.

Autor, D., Dorn, D., Hanson, G., Majlesi, K., 2020. Importing political polarization? the electoral consequences of rising trade exposure. American Economic Review 110, 3139–83.

Autor, D.H., Dorn, D., Hanson, G.H., 2013. The china syndrome: Local labor market effects of import competition in the united states. American Economic Review 103, 2121–68.

Autor, D.H., Dorn, D., Hanson, G.H., 2016. The china shock: Learning from labor-market adjustment to large changes in trade. Annual Review of Economics 8, 205–240.

Baldwin, R.E., Magee, C.S., 2000. Is trade policy for sale? congressional voting on recent trade bills. Public Choice 105, 79–101.

Bartels, L.M., 1993. Messages received: The political impact of media exposure. American political science review 87, 267–285.

Bateman, D.A., Clinton, J., Lapinski, J.S., 2017. A house divided? political conflict and polarization in the u.s. congress, 1877- 2011. American Journal of Political Science 61, 698–714.

Blanchard, E.J., Bown, C.P., Chor, D., 2019. Did Trump's Trade War Impact the 2018 Election? Technical Report. National Bureau of Economic Research.

Breinlich, H., 2014. Heterogeneous firm-level responses to trade liberalization: A test using stock price reactions. Journal of International Economics 93, 270–285.

Bugni, F.A., Canay, I.A., Shi, X., 2015. Specification tests for partially identified models defined by moment inequalities. Journal of Econometrics 185, 259–282.

Canen, N., Jackson, M.O., Trebbi, F., 2019. Endogenous networks and legislative activity. Mimeo .

Canen, N., Kendall, C., Trebbi, F., 2020. Unbundling ization. Econometrica 88, 1197–1233.

Che, Y., Lu, Y., Pierce, J.R., Schott, P.K., Tao, Z., 2020. Does trade liberalization with China influence US elections? Technical Report. National Bureau of Economic Research.

Clinton, J., Jackman, S., Rivers, D., 2004. The statistical analysis of roll call data. American Political Science Review 98, 355–370.

Colantone, I., Stanig, P., 2018. The trade origins of economic nationalism: Import competition and voting behavior in western europe. American Journal of Political Science 62, 936–953.

Cruz, C., Keefer, P., Labonne, J., Trebbi, F., 2018. Making policies matter: Voter responses to campaign promises. Technical Report. National Bureau of Economic Research.

Dickstein, M.J., Morales, E., 2018. What do exporters know? The Quarterly Journal of Economics 133, 1753–1801.

Drazen, A., Masson, P.R., 1994. Credibility of policies versus credibility of policymakers. The Quarterly Journal of Economics 109, 735–754.

Egorov, G., 2015. Single-issue campaigns and multidimensional politics. Technical Report. National Bureau of Economic Research.

Fajgelbaum, P.D., Goldberg, P.K., Kennedy, P.J., Khandelwal, A.K., 2020. The return to protectionism. The Quarterly Journal of Economics 135, 1–55.

Feenstra, R.C., Ma, H., Xu, Y., 2019. Us exports and employment. Journal of International Economics 120, 46–58.

Feigenbaum, J.J., Hall, A.B., 2015. How legislators respond to localized economic shocks: Evidence from chinese import competition. The Journal of Politics 77, 1012–1030.

Feler, L., Senses, M.Z., 2017. Trade shocks and the provision of local public goods. American Economic Journal: Economic Policy 9, 101–43.

Fenno, R.F., 1978. Home Style: House Members in their Districts. Boston: Little, Brown.

Gentzkow, M., Shapiro, J.M., Taddy, M., 2019. Measuring group differences in high-dimensional choices: method and application to congressional speech. Econometrica 87, 1307–1340.

Greenland, A., Ion, M., Lopresti, J., Schott, P., 2020. Using equity market reactions to infer exposure to trade liberalization .

Greenland, A., Lopresti, J., 2016. Import exposure and human capital adjustment: Evidence from the us. Journal of International economics 100, 50–60.

Greenland, A., Lopresti, J., McHenry, P., 2019. Import competition and internal migration. Review of Economics and Statistics 101, 44–59.

Heckman, J.J., Snyder, J.M., 1997. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. The RAND Journal of Economics 28.

Ho, K., Pakes, A., 2014. Hospital choices, hospital prices, and financial incentives to physicians. American Economic Review 104, 3841–84.

Jenkins, J.A., 2000. Examining the robustness of ideological voting: evidence from the confederate house of representatives. American Journal of Political Science , 811–822.

Johnson, R.D., 2005. Congress and the cold war. Cambridge University Press.

Jones, B.D., Baumgartner, F.R., 2005. The politics of attention: How government prioritizes problems. University of Chicago Press.

Kalt, J.P., Zupan, M.A., 1984. Capture and ideology in the economic theory of politics. The American Economic Review 74, 279–300.

Kalt, J.P., Zupan, M.A., 1990. The apparent ideological behavior of legislators: Testing for principal-agent slack in political institutions. The Journal of Law and Economics 33, 103–131.

Kendall, C., Nannicini, T., Trebbi, F., 2015. How do voters respond to information? evidence from a randomized campaign. American Economic Review 105, 322–53.

Krehbiel, K., 1992. Information and legislative organization. University of Michigan Press.

Ladewig, J.W., 2010. Ideological polarization and the vanishing of marginals: Retrospective roll-call voting in the us congress. The Journal of Politics 72, 499–512.

Le Cam, L., 1960. An approximation theorem for the poisson binomial distribution. Pacific Journal of Mathematics 10, 1181–1197.

Lee, D., Moretti, E., Butler, M., 2004. Do voters affect or elect policies? evidence from the u.s. house. Quarterly Journal of Economics 119, 807–860.

Levitt, S.D., 1996. How do senators vote? disentangling the role of voter preferences, party affiliation, and senator ideology. The American Economic Review 86, 425–441.

Manski, C., 1991. Nonparametric estimation of expectations in the analysis of discrete choice under uncertainty, in: nonparametric and Semiparametric methods in Econometrics and Statistics: proceedings of the fifth International Symposium in Economic Theory and Econometrics, Cambridge University Press. p. 259.

McCarty, N., Poole, K.T., Rosenthal, H., 2006. Polarized America: The Dance of Ideology and Unequal Riches. Cambridge: MIT Press.

Mian, A., Sufi, A., Trebbi, F., 2010. The political economy of the us mortgage default crisis. American Economic Review 100.

Mian, A., Sufi, A., Trebbi, F., 2014. Resolving debt overhang: Political constraints in the aftermath of financial crises. American Economic Journal: Macroeconomics 6, 1–28.

Nokken, T.P., Poole, K.T., 2004. Congressional party defection in american history. Legislative Studies Quarterly 29, 545–568.

Peltzman, S., 1984. Constituent interest and congressional voting. The Journal of Law and Economics 27, 181–210.

Pierce, J.R., Schott, P.K., 2016. The surprisingly swift decline of us manufacturing employment. American Economic Review 106, 1632–62.

Pierce, J.R., Schott, P.K., 2020. Trade liberalization and mortality: evidence from us counties. American Economic Review: Insights 2, 47–64.

Poole, K., Rosenthal, H., 2017. Ideology and Congress: A Political Economic History of Roll Call Voting. 2nd Edition Routledge.

Poole, K.T., 2007. Changing minds? not in congress! Public Choice 131, 435–451.

Poole, K.T., Rosenthal, H., 1984. The polarization of american politics. Journal of Politics 46, 1061–1079.

Poole, K.T., Rosenthal, H., 1997. Congress: A Political-Economic History of Roll Call Voting. New York: Oxford University Press.

Pregelj, V.N., 1998. Most-Favored-Nation Status: Congressional Consideration, 1989-1998. Technical Report. Congressional Research Service.

Pregelj, V.N., 2001. Most-Favored-Nation Status of the People's Republic of China. Technical Report. Congressional Research Service.

Primiceri, G.E., 2006. Why inflation rose and fell: policy-makers' beliefs and us postwar stabilization policy. The Quarterly Journal of Economics 121, 867–901.

Rodrik, D., 1995. Political economy of trade policy. Handbook of international economics 3, 1457–1494.

Sargent, T., Williams, N., Zha, T., 2006. Shocks and government beliefs: The rise and fall of american inflation. American Economic Review 96, 1193–1224.

Stimson, J.A., MacKuen, M.B., Erikson, R.S., 1995. Dynamic representation. American political science review , 543–565.

Stratmann, T., 2005. Some talk: Money in politics. a (partial) review of the literature, in: Policy challenges and political responses. Springer, pp. 135–156.

Yatchew, A., Griliches, Z., 1985. Specification error in probit models. The Review of Economics and Statistics , 134–139.

Figure 1: Roll call votes Pro China by Party
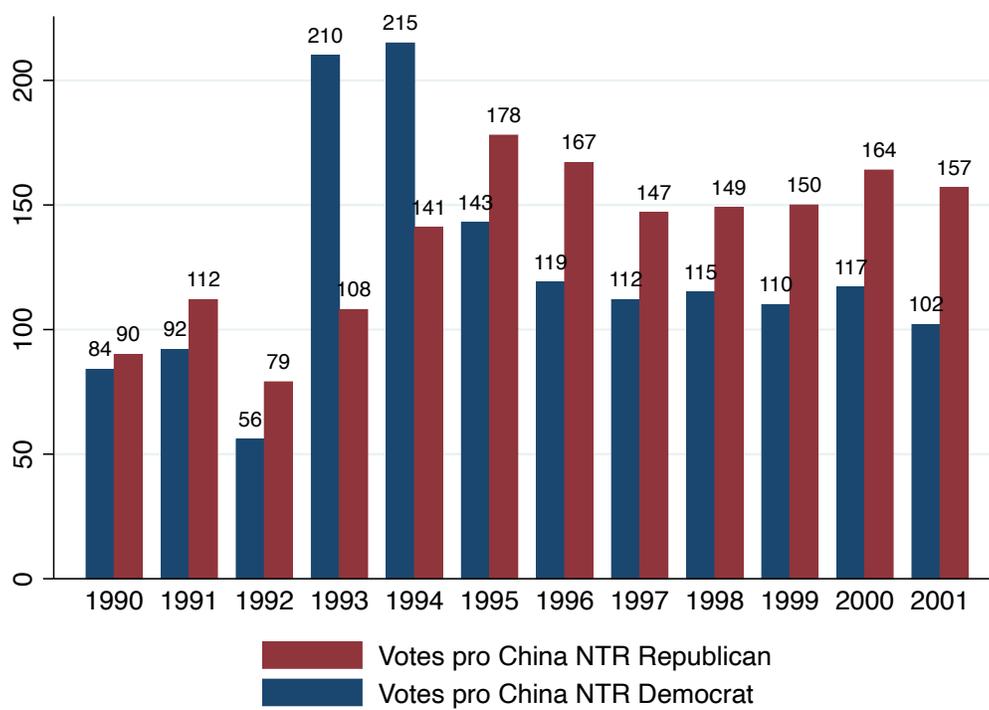
Figure 2: Roll call vote switching - Democrats
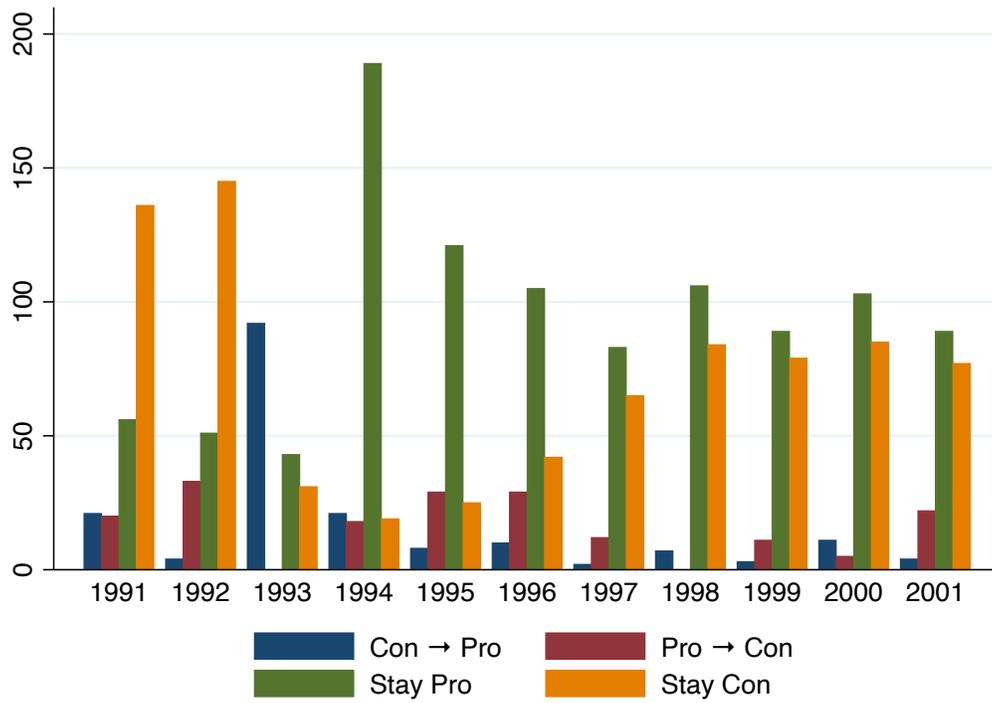
Figure 3: Roll call vote switching - Republicans
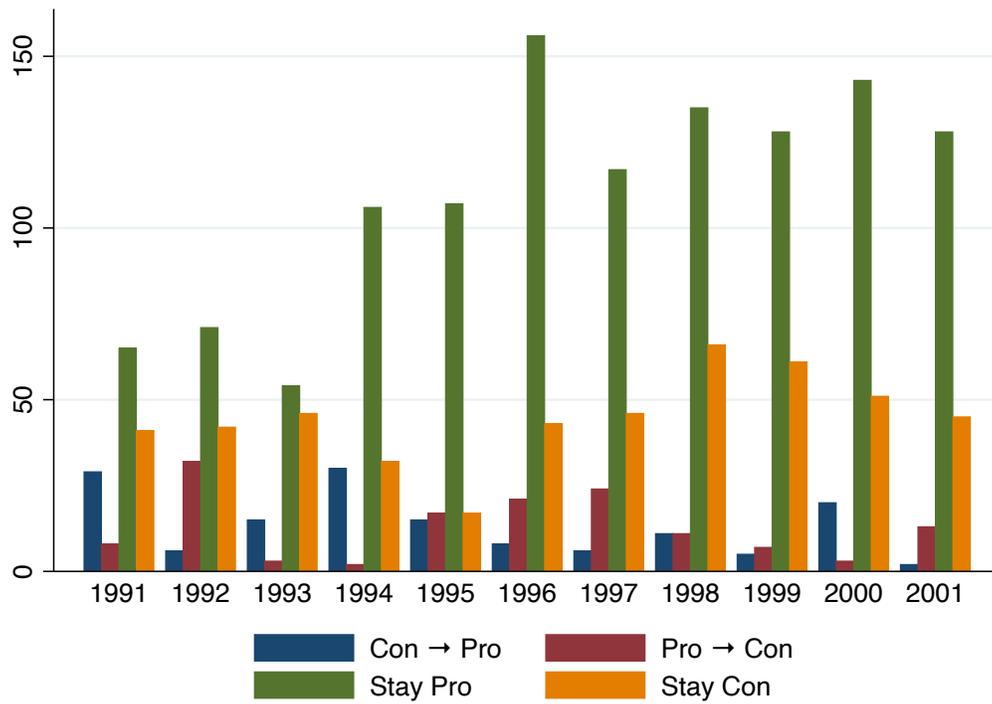
Figure 4: Import Shocks from China and Pro-China NTR Vote Share
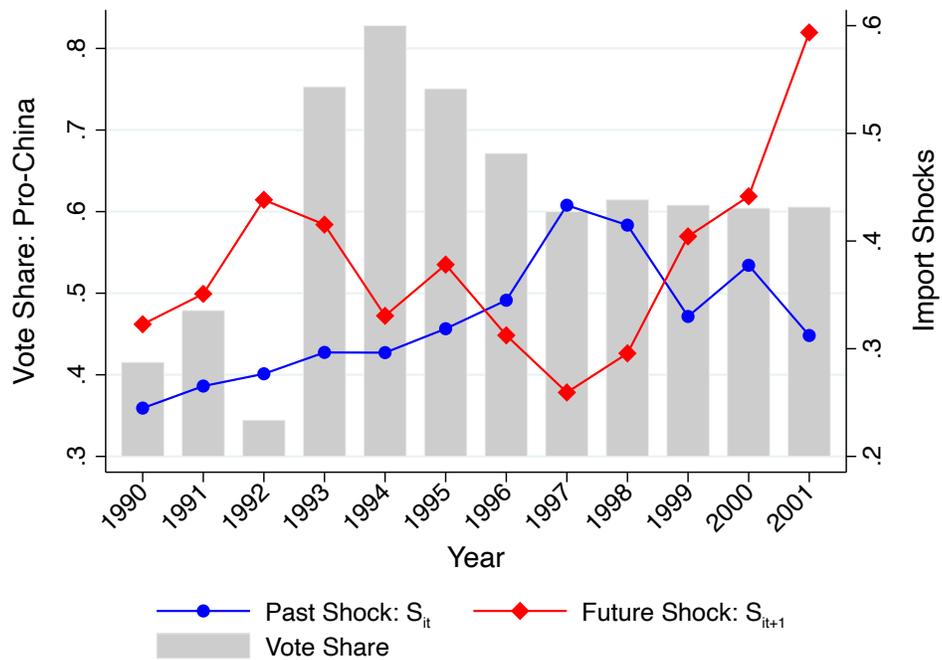
Notes: The bar chart shows the share of votes in favor of renewing China's NTR status for the bills introduced in House over 1990 to 2001. The data in 2000 also includes the bill HR 4444 which granted China permanent normal trade relations. The dot blue line represents the average of past import shock across congressional districts, and the diamond red line shows the average of the future import shock. 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.

Figure 5: The China Shock over Time
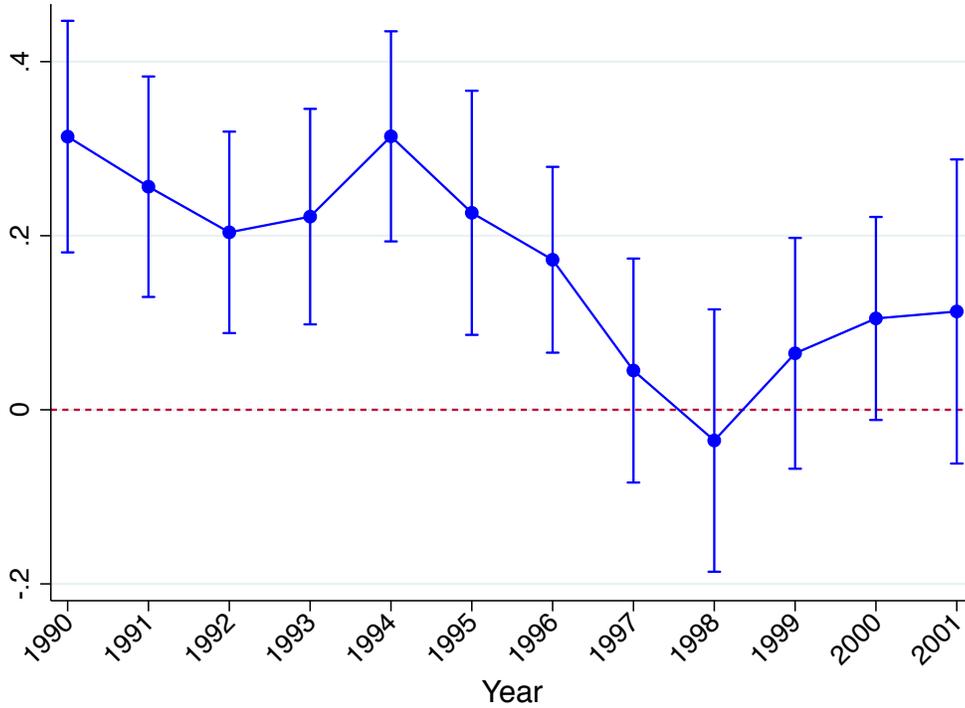
Notes: We explore the autocorrelation of import supply shock from China at the 4-digit SIC level by estimating the following equation:

$$\Delta \ln M^{oth}_{k,t,t+5} = \alpha_t \Delta \ln M^{oth}_{k,t-5,t} + D_t + \varepsilon_{kt}$$

where $\Delta \ln M^{oth}_{k,t-5,t}$ (respectively, $\Delta \ln M^{oth}_{k,t,t+5}$) measures the change in log import from China by eight developed countries (Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland) over the period $t-5$ to t (respectively, t to $t+5$). D_t is the year fixed effects, and α_t captures the autocorrelation of import supply shock from China in different periods. Standard errors are clustered at the 4-digit SIC level. We use the Davis-Haltiwanger-Schuh approximation of the log growth rate, i.e., $\Delta \ln M^{oth}_{k,t-5,t} \approx 2 \left(\frac{M^{oth}_{k,t} - M^{oth}_{k,t-5}}{M^{oth}_{k,t} + M^{oth}_{k,t-5}} \right)$ and $\Delta \ln M^{oth}_{k,t,t+5} \approx 2 \left(\frac{M^{oth}_{k,t+5} - M^{oth}_{k,t}}{M^{oth}_{k,t+5} + M^{oth}_{k,t}} \right)$, to avoid dropping observations where imports are zero. The point estimates of α_t and their 95% confidence intervals are reported in the figure.

Figure 6: An Illustrative Example: Effects of Improving Information

A

B

∫Φ(x+η)dG(η)<Φ(x) when x>0

∫Φ(x+η)dG(η)>Φ(x) when x<0

D

C

0
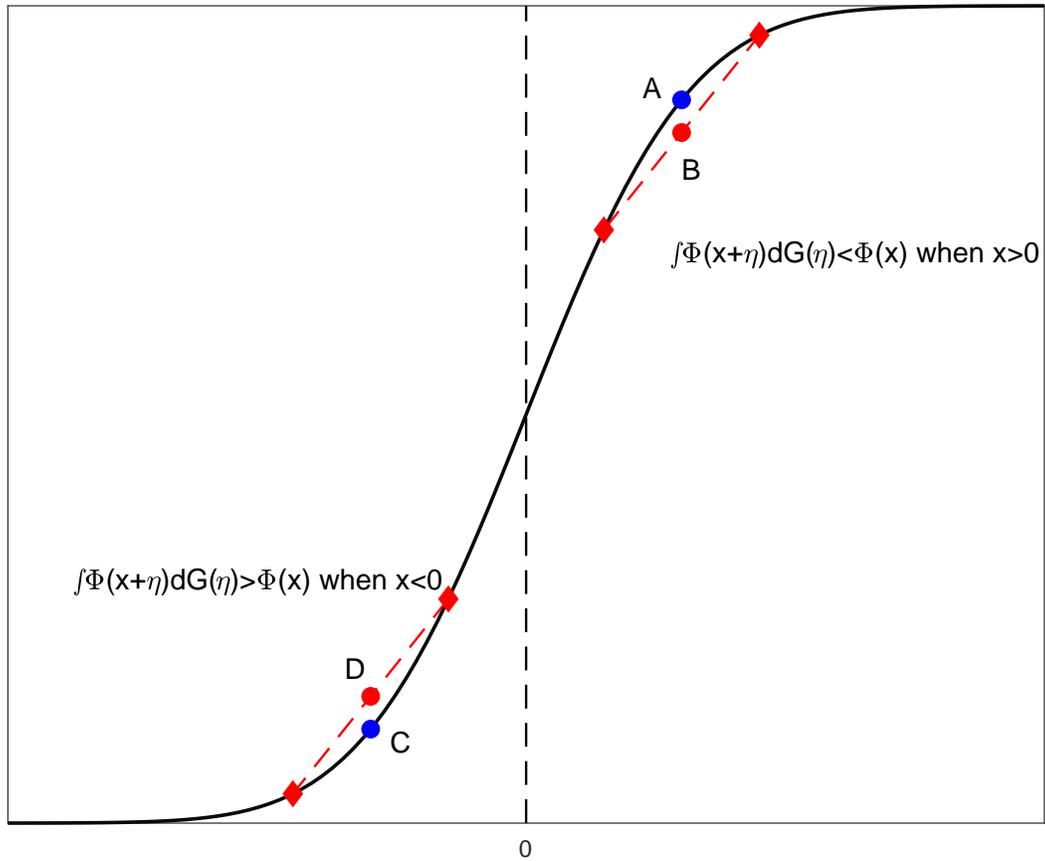
Notes: The figure presents an illustrative example on how the provision of information on the China shock changes the pro-China vote share. The black solid curve represent the standard normal cumulative density function. In this figure x corresponds to the component $a_t\theta_i + b_t + \delta_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b]$ and η corresponds to the component $\delta_t\varepsilon_{i,t+1}$. In this example, the distribution of η has a mean zero and is symmetric. The spread represented by the red dashed lines is determined by the variation of η.

Table 1: MLE Estimates Based on Different Information Sets

	$\hat{a}$ (se.)	$\hat{b}$ (se.)	$\hat{\delta}$ (se.)
Panel A: 1997-2001			
Minimal information $Z_{i,t}=\{ShareMfg_{i,t},\theta_i\}$	0.650 (0.063)	0.262 (0.025)	-0.810 (0.314)
Baseline information $Z_{i,t}=\{S_{i,t}, ShareMfg_{i,t},\theta_i\}$	0.643 (0.062)	0.263 (0.025)	-0.786 (0.279)
Perfect foresight	0.619 (0.062)	0.263 (0.025)	-0.018 (0.211)
Panel B: 1993-1996			
Minimal information $Z_{i,t}=\{ShareMfg_{i,t},\theta_i\}$	-0.046 (0.084)	0.678 (0.033)	-0.440 (0.611)
Baseline information $Z_{i,t}=\{S_{i,t}, ShareMfg_{i,t},\theta_i\}$	-0.043 (0.084)	0.679 (0.033)	-1.116 (0.549)
Perfect foresight	-0.051 (0.083)	0.678 (0.033)	-0.517 (0.452)
Panel C: 1990-1992			
Minimal information $Z_{i,t}=\{ShareMfg_{i,t},\theta_i\}$	1.019 (0.104)	-0.179 (0.037)	0.065 (0.351)
Baseline information $Z_{i,t}=\{S_{i,t}, ShareMfg_{i,t},\theta_i\}$	1.000 (0.105)	-0.181 (0.037)	-0.981 (0.288)
Perfect foresight	1.013 (0.104)	-0.180 (0.037)	-0.669 (0.244)
Panel D: 1990-2001			
Minimal information $Z_{i,t}=\{ShareMfg_{i,t},\theta_i\}$	0.532 (0.044)	0.280 (0.017)	-0.453 (0.222)
Baseline information $Z_{i,t}=\{S_{i,t}, ShareMfg_{i,t},\theta_i\}$	0.534 (0.044)	0.280 (0.017)	-0.856 (0.198)
Perfect foresight	0.523 (0.044)	0.280 (0.017)	-0.333 (0.147)

Notes: This table reports the MLE estimates based on equation (6). For the case of minimal information, we replace the term $\mathbb{E}\left[S_{i,t+1}|Z_{i,t}\right]$ by the predicted value of the OLS regression: $S_{i,t+1} = \beta_0 + \beta_1\theta_i + \beta_2 ShareMfg_{i,t} + \varepsilon_{i,t+1}$. For the case of baseline information, the term is replaced by the predicted value of the OLS regression: $S_{i,t+1} = \beta_0 + \beta_1\theta_i + \beta_2 ShareMfg_{i,t} + \beta_3 S_{i,t} + \varepsilon_{i,t+1}$. For the case of perfect foresight, the term is replaced by $S_{i,t+1}$. Robust standard errors are reported in parentheses.

Table 2: Relevance of Manufacturing Employment Share and Past Import Shock in Explaining Future Import Shocks

Sample Period: Dependent Variable: $S_{i,t+1}$	1990-2001 (1)	1997-2001 (2)	1993-1996 (3)	1990-1992 (4)
$S_{i,t}$	0.404***	0.731***	0.214***	0.654***
	(0.017)	(0.041)	(0.013)	(0.033)
$ShareMfg_{i,t}$	0.705***	0.680***	0.457***	0.700***
	(0.021)	(0.036)	(0.020)	(0.045)
Observations	5,494	2,564	1,698	1,232
R^2	0.556	0.584	0.674	0.695

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. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Parameter Confidence Sets and Specification Test p-values

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 $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[0.450, 0.720]	[0.165, 0.260]	[-1.975, -0.225]	0.405	0.405	0.405	2564
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[0.435, 0.750]	[0.185, 0.290]	[-1.560, -0.075]	0.290	0.275	0.275	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 $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[-0.280, 0.100]	[0.583, 0.703]	[-2.300, 0.900]	0.330	0.330	0.330	1698
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[-0.325, 0.130]	[0.598, 0.740]	[-3.125, -0.125]	0.395	0.395	0.395	1698
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1}|S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	–	–	–	0.040	0.035	0.035	1698
1990-1992	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[0.800, 1.550]	[-0.325, -0.125]	[-1.125, 2.125]	0.955	0.955	0.955	1232
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[1.025, 1.438]	[-0.275, -0.150]	[-1.300, 0.000]	0.165	0.145	0.145	1232
	Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, S_{i,t+1} - E[S_{i,t+1}|S_{i,t}, ShareMfg_{i,t}, \theta_i], \theta_i\}$						
	[1.000, 1.550]	[-0.270, -0.128]	[-1.624, 0.096]	0.255	0.240	0.240	1232
1990-2001	**Panel A:** Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$						
	[0.463, 0.625]	[0.210, 0.270]	[-1.163, -0.150]	0.185	0.185	0.185	5494
	Panel B: Baseline information $Z_{i,t} = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$						
	[0.463, 0.650]	[0.235, 0.287]	[-1.438, -0.500]	0.190	0.190	0.190	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 4: Parameter Confidence Sets and Specification Test p-values: Heterogeneity by Party
(Baseline 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	[2.225, 4.625]	[0.850, 1.800]	[-4.425, -0.015]	0.315	0.265	0.265	1229
	Republicans	–	–	–	0.010	0.010	0.010	1326
1993-1996	Democracts	[0.800, 3.400]	[1.050, 2.050]	[-6.325, -0.025]	0.380	0.380	0.380	888
	Republicans	–	–	–	0.020	0.015	0.015	806
1990-1992	Democracts	[1.875, 4.375]	[0.050, 0.850]	[-3.700, -0.400]	0.415	0.410	0.410	745
	Republicans	[-1.300, 1.400]	[-0.300, 0.650]	[-1.500, 1.800]	0.315	0.310	0.310	485
1990-2001	Democracts	[1.375, 3.325]	[0.570, 1.290]	[-4.412, -0.417]	0.990	0.990	0.990	2862
	Republicans	–	–	–	0.010	0.010	0.010	2617

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

Table 5: Parameter Confidence Sets and Specification Test p-values: Heterogeneity by Win Margin
(Baseline 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	Win-margin>median	[0.600, 1.225]	[0.060, 0.270]	[-1.375, 1.500]	0.890	0.890	0.890	1205
	Win-margin<median	[-0.125, 0.500]	[0.250, 0.438]	[-3.413, -0.525]	0.680	0.675	0.675	1207
1993-1996	Win-margin>median	[-0.375, 0.350]	[0.475, 0.700]	[-3.250, 2.750]	0.560	0.560	0.560	799
	Win-margin<median	[-1.050, 0.300]	[0.625, 0.975]	[-8.550, -0.225]	0.965	0.965	0.965	800
1990-1992	Win-margin>median	[0.875, 2.125]	[-0.465, -0.105]	[-2.875, 3.312]	0.765	0.755	0.755	568
	Win-margin<median	[0.450, 1.700]	[-0.325, -0.025]	[-4.375, -0.750]	0.510	0.465	0.465	569
1990-2001	Win-margin>median	[0.550, 0.888]	[0.120, 0.232]	[-0.900, 0.800]	0.450	0.450	0.450	2573
	Win-margin<median	[0.013, 0.475]	[0.282, 0.410]	[-2.565, -0.825]	0.615	0.615	0.615	2575

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

Table 6: Parameter Confidence Sets and Specification Test p-values: Heterogeneity by Tenure
(Baseline 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	Tenure>median	[0.750, 1.075]	[0.158, 0.255]	[-0.850, 0.600]	0.120	0.105	0.105	1372
	Tenure<median	[0.180, 0.700]	[0.238, 0.375]	[-2.750, -0.875]	0.185	0.165	0.165	1176
1993-1996	Tenure>median	[-0.400, 0.350]	[0.525, 0.700]	[-3.350, 1.363]	0.460	0.460	0.460	863
	Tenure<median	[-0.700, 0.237]	[0.650, 0.925]	[-5.875, 0.312]	0.630	0.615	0.615	821
1990-1992	Tenure>median	[0.825, 2.400]	[-0.175, 0.175]	[-5.513, 0.175]	0.920	0.920	0.920	572
	Tenure<median	[0.938, 1.750]	[-0.495, -0.270]	[-1.600, 0.900]	0.260	0.235	0.235	634
1990-2001	Tenure>median	[0.615, 0.965]	[0.200, 0.287]	[-1.758, -0.270]	0.275	0.275	0.275	2488
	Tenure<median	[0.238, 0.475]	[0.262, 0.338]	[-1.502, -0.355]	0.145	0.135	0.135	2950

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

Table 7: Parameter Confidence Sets and Specification Test p-values: Heterogeneity by Campaign Contributions from Business
(Baseline 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	Money>median	[-0.350, 0.450]	[0.492, 0.720]	[-2.850, -0.600]	0.360	0.360	0.360	1273
	Money<median	[0.525, 1.078]	[-0.045, 0.135]	[-2.200, 0.425]	0.380	0.345	0.345	1273
1993-1996	Money>median	[-1.000, 0.250]	[0.725, 1.075]	[-7.000, -0.400]	0.715	0.695	0.695	844
	Money<median	[-0.312, 0.588]	[0.450, 0.700]	[-3.650, 2.125]	0.665	0.665	0.665	844
1990-1992	Money>median	[0.150, 1.100]	[-0.185, 0.025]	[-2.038, 0.325]	0.370	0.370	0.370	613
	Money<median	[1.112, 3.475]	[-0.400, 0.275]	[-7.325, 0.775]	0.830	0.830	0.830	616
1990-2001	Money>median	[0.550, 0.438]	[0.390, 0.490]	[-2.090, -0.900]	0.220	0.215	0.215	2730
	Money< median	[0.510, 0.912]	[0.100, 0.220]	[-1.770, 0.128]	0.440	0.440	0.440	2733

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

Table 8: Parameter Confidence Sets and Specification Test p-values:
Votes on NTR with Vietnam (1998-2001)

CS of a	CS of b	CS of δ	p-value BP	p-value RC	p-value RS	Num obs.
		Panel A: Minimal information $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$				
[-1.300, -0.300]	[0.600, 0.938]	[-85.000, -21.625]	0.990	0.990	0.990	1595
		Panel B: Baseline information $Z_{i,t} = \{S_{i,t}^V, ShareMfg_{i,t}, \theta_i\}$				
[-1.250, -0.613]	[0.630, 0.792]	[-74.075, -6.700]	0.435	0.435	0.435	1595
		Panel C: Perfect foresight $Z_{i,t} = \{S_{i,t}^V, ShareMfg_{i,t}, S_{i,t+1}^V - E[S_{i,t+1}^V|S_{i,t}^V, ShareMfg_{i,t}, \theta_i], \theta_i\}$				
–	–	–	0.025	0.020	0.020	1595

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

Table 9: Effects of Improving Information

	1997-2001	1993-1996	1990-1992
(1) Change in share of votes pro-CHN (%)	[-0.030, 0.012]	[-0.161, -0.000]	[-0.008, 0.064]
(2) Share of always pro-CHN (%)	[55.956, 61.668]	[70.387, 76.966]	[36.629, 42.352]
(3) Share of pro-CHN to against-CHN (%)	[0.078, 1.695]	[0.054, 1.606]	[0.000, 1.272]
(4) Share of against-CHN to pro-CHN (%)	[0.077, 1.694]	[0.054, 1.482]	[0.000, 1.316]
(5) Share of always against-CHN (%)	[36.723, 42.187]	[21.998, 27.645]	[56.266, 62.309]

Notes: The simulation in this table is based on the assumption that the true information set is the Baseline information set $\mathcal{I}_{i,t}^b = \{S_{i,t}, ShareMfg_{i,t}, \theta_i\}$.

Table 10: Effects of Heightening Constituent Interests

	1997-2001	1993-1996	1990-1992
	Panel A: Baseline		
(1) Value of δ	[-1.560, -0.075]	[-3.125, -0.125]	[-1.300, 0.000]
(2) Share of votes pro-CHN (%)	[57.663, 61.760]	[71.999, 77.034]	[37.682, 42.427]
	Panel B: Lower bound of CS for Democrats		
(3) Value of δ	-4.425	-6.325	-3.700
(4) Share of votes pro-CHN (%)	[17.485, 33.071]	[35.132, 58.356]	[11.983, 23.131]
	Panel C: Lower bound of CS for Win Margin < median		
(5) Value of δ	-3.413	-8.550	-4.375
(6) Share of votes pro-CHN (%)	[24.372, 42.458]	[24.606, 45.485]	[9.682, 19.424]

Notes: The simulation in this table is based on the assumption that the true information set is the Baseline information set $\mathcal{I}_{i,t}^{b} = \{S_{i,t}, ShareMfg_{i,t}, \theta_{i}\}$.

Online Appendix to

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

Matilde Bombardini, Bingjing Li, Francesco Trebbi

- Not for Publication -

A Empirical Model: Additional Details

A.1 Microfoundation of Equation (1)

The expression used in equation (1) is a common representation of expected electoral support (Bartels, 1993; Canen et al., 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 et al., 2001; Lee et al., 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 et al. (2015).

The probability voter j votes for i is then:

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

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 $\mathbf{1}\{d_{i,t-1} = \text{vote for } q_{t-1}\}$, and votes for i' if instead $\mathbf{1}\{d_{i,t-1} = \text{vote for } x_{t-1}\}$, so that:

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

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:

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

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 $t+1$:

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

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 $\gamma_1 - \gamma_2 < 0$.

A.1.1 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. 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:

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

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 $\mathbf{X}_{i,t}$ 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, 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. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

A.2 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),[40] 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

---

[40]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 $S_{k,t+1}$.

at the broader sector level that 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
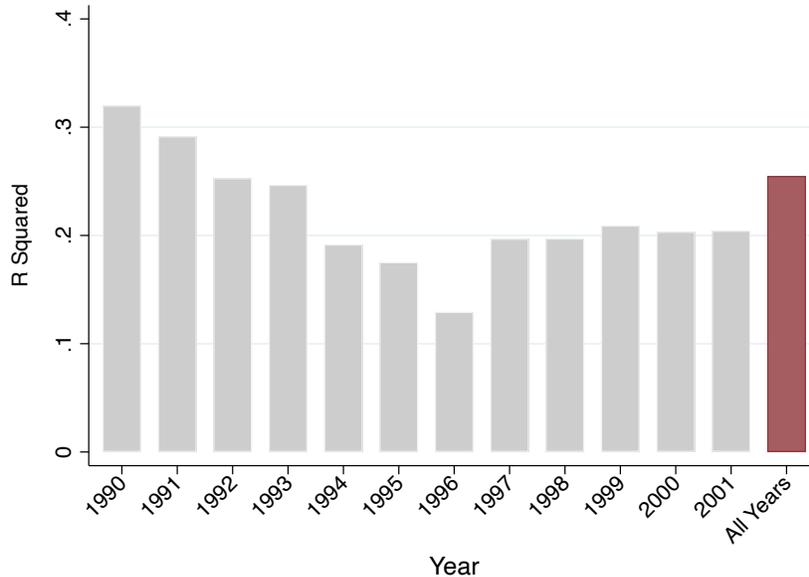
2. The discussion so far has relied on the assumption that agents perfectly know the weights $w_{ik,t}$, which is a stringent requirement. While it is plausible that politicians know the manufacturing employment share in their districts, having the exact knowledge of current employment share for each 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

54

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

A.3 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

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

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:

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

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.[41] 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.

---

[41]See Jones and Baumgartner (2005).

B 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 (6). 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 $\mathcal{I}_{i,t}^b = \{S_{it}, ShareMfg_{i,t}, \theta_i\}$. More specifically, $S_{i,t+1}$ is simulated according to the following linear model:

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

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 $\sigma_\varepsilon = 0.525$.[42] 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 $Y_{i,t} = \mathbf{1}\{a\theta_i + b + \delta\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b] - \xi_{i,t} > 0\}$ 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 $\mathcal{I}_{i,t}^m$ or perfect foresight $\mathcal{I}_{i,t}^p$. For the case of minimal information set, we estimate ω_t by maximizing the following log-likelihood:

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

where the expectation is estimated as:

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

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

---

[42]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.

average parameter estimates are very close to the 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:

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

where the error term becomes $\rho_{i,t} = \delta_t \left\{\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}^b\right] - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}^m\right]\right\} - \xi_{i,t}$, with a standard deviation of $\sigma_\rho > \sigma_\xi = 1$. It is straightforward to show that $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}^b\right] - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}^m\right]$ and $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}^m\right]$ 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)

C Estimation strategy, specification tests and counterfactual simulations: details

C.1 Estimation implementation

Conditional moments (9), (10), (13) and (14) 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.[43] 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.[44]

We collect the four moment inequalities (9), (10), (13) and (14) 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:

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

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 consititued by θ_i and $ShareMfg_{i,t}$. For the minimal information case, we have $8 \times 4 = 32$ moment inequalities, which

---

[43]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.

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

we use to construct confidence sets, as explained in section 5.[45] The choice of the instrument functions does 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.

C.2 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

---

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

the null hypothesis $H_o : \omega^* = \omega_p$. Here, ω^* denotes the true parameter vector. More specifically, the simulation is based on R draws 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:

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

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).[46] 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 $\hat{\Omega}^{1-\alpha}$.[47] In the main text, we set $\alpha = 0.05$, and report the 95% confidence sets.

C.3   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 $Z_{i,t} \nsubseteq \mathcal{I}_{i,t}$, 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 et al. (2015).

C.3.1   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.[48] As pointed out by Bugni et al. (2015), this test is too conservative. Therefore, we turn to the test RC and the test RS proposed by the authors as follows.

---

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

[47]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 5.4 and Appendix E.3, we fill the 4-dimensional space with 160,000 equidistant grids.

[48]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).

C.3.2 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 $\Omega_g^* = \arg\min_{\Omega_p}(Q(\omega_p))$.[49] 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 $T \le \hat{c}^{RC}(1-\lambda^{RC})$. Then, λ^{RC} is the corresponding p-value.

C.3.3 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 $T \le \hat{c}^{RS}(1-\lambda^{RS})$. 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, $\hat{c}^{RC}(1-\lambda) \ge \hat{c}^{RS}(1-\lambda)$, and the p-value of the test RS will be no larger than that of the test RC.

C.4 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:

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

---

[49] 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:

$$N^{++}(\omega_t, \mathcal{I}_{i,t}^b \to \mathcal{I}_{i,t}^p, \mathcal{N}_t)$$

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

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

$$N^{-+}(\omega_t, \mathcal{I}_{i,t}^b \to \mathcal{I}_{i,t}^p, \mathcal{N}_t)$$

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

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

$$N^{--}(\omega_t, \mathcal{I}_{i,t}^b \to \mathcal{I}_{i,t}^p, \mathcal{N}_t)$$

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

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

D Data appendix

D.1 Trade, employment and output data

This subsection describes the data sources that we employ to construct the commuting-zone-level measures (15) and (16). 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. We concord these data to four-digit Standard Industry Classification (SIC) industries.[50] Following Autor et al. (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. 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. 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 et al. (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 et al. (2013).[51]

D.2 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 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

---

[50]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 et al. (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.

[51]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 et al. (2013), we concord the post-1997 data to the four-digit SIC industries.

import supply shocks from Vietnam at the CZ level according to

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

where $\Delta M_{k,t+1}^{oth,V}$ 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

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

where $\Delta M_{k,t}^{oth,V}$ 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 4.2. The CD-level measures are denoted by $S_{i,t+1}^V$ and $S_{i,t}^V$. The magnitude of the import shock from Vietnam is several orders smaller than that from China. The standard deviation of $S_{i,t+1}^V$ 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) $S_{i,t+1}^V$ on (de-trended) $S_{i,t}^V$ 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.

E   Additional results

E.1   Evidence from congressional speech

The data on congressional speech is obtained from (Gentzkow et al., 2019). 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.[52] The number of relevant speeches is related to import shocks from China according to:

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

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 5 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 China.

---

[52]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.

E.2 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

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

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 2.

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 $\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right] = \gamma_t \mathbb{E}\left[S_{i,t+1}^{NTR} - S_{i,t+1}^{NNTR}|\mathcal{I}_{i,t}\right]$. For the estimation, we use $S_{i,t+1} = S_{i,t+1}^{NTR} - S_{i,t+1}^{NNTR}$ 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, $S_{j,t+1} = S_{j,t+1}^{NTR} - S_{j,t+1}^{NNTR}$ (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_{cuk}^{NNTR} and τ_{cuk}^{NTR}, and $\tau_{cuk}^{NNTR} > \tau_{cuk}^{NTR}$. The corresponding shocks are denoted by S_k^{NNTR} and S_k^{NTR}. Hence,

$$\begin{aligned}
S_k^{NTR} - S_k^{NNTR} &= S_k^{NTR} \times \left[1 - \left(\frac{1 + \tau_{cuk}^{NNTR}}{1 + \tau_{cuk}^{NTR}}\right)^{-\rho}\right] \\
&\approx \rho S_k^{NTR} \times \left(\ln(1 + \tau_{cuk}^{NNTR}) - \ln(1 + \tau_{cuk}^{NTR})\right),
\end{aligned} \tag{E.2}$$

where the approximation follows from the fact that $\ln(1-x) \approx -x$ when x has a small and positive value.[53] Note that S_k^{NTR} is the observed China supply shock. It is proxied by the surge in import penetration from China in eight other developed countries in Autor, Dorn and Hanson (2013), i.e., $\left(\frac{\Delta M_{k,t+1}^{oth}}{Y_{k,t} + M_{k,t} - X_{k,t}}\right)$.

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

---

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

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

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: $\sum_k \frac{L_{jk,t}}{L_{j,t}} (S_{k,t+1}^{NTR} - S_{k,t+1}^{NNTR})$. The term $NTRGap_{k,t} \times \frac{\Delta M_{k,t+1}^{oth}}{Y_{k,t} + M_{k,t} - X_{k,t}}$ 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:

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

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.

E.3   Introducing export shocks

Similar to Feenstra et al. (2019), we construct the future export shock from China in CZ j according to:

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

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 $Z_{i,t} = \{S_{i,t}, S_{i,t}^X, ShareMfg_{i,t}, \theta_i\}$. 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.

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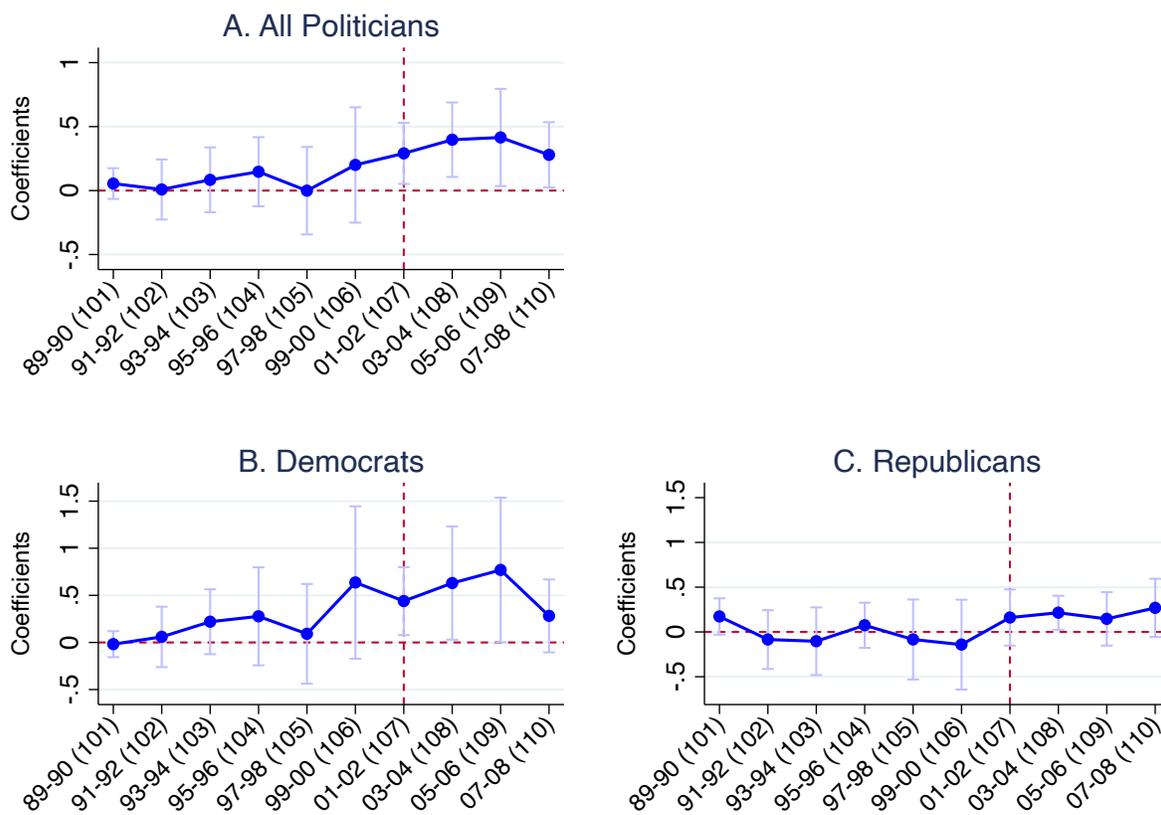
Notes: This figure plots the estimated coefficients β_c in equation (E.1). Error bands show 90% confidence intervals. Standard errors are clustered at the congressional district level. The vertical red dashed line indicates the time of China's accession to the WTO.

Figure E.2: Distribution of Expectational Errors
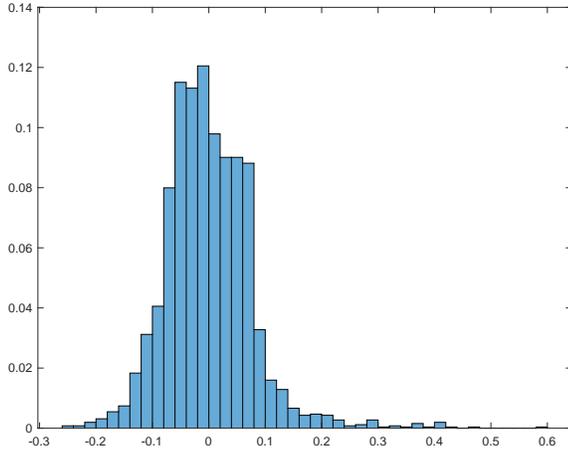
A. 1997-2001
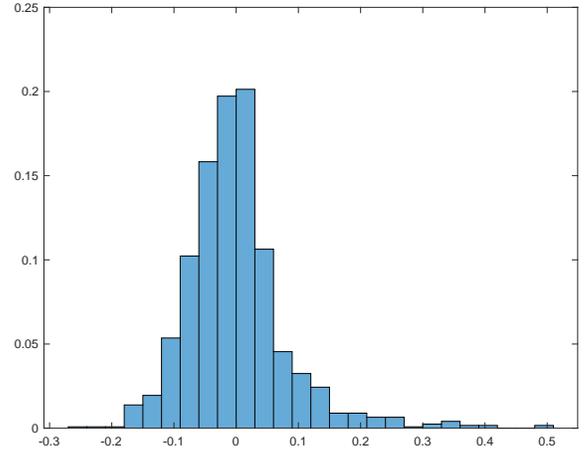
B. 1993-1996
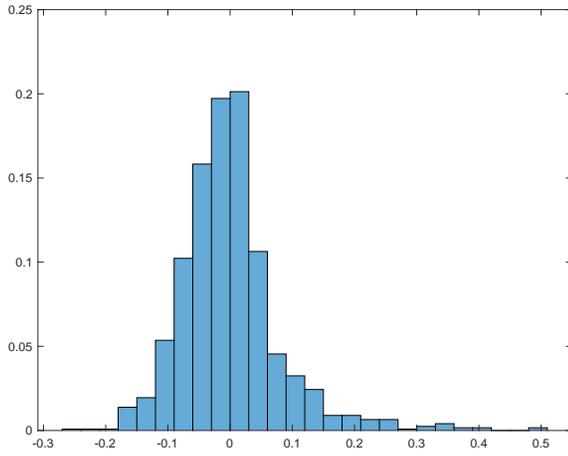
C. 1990-1992
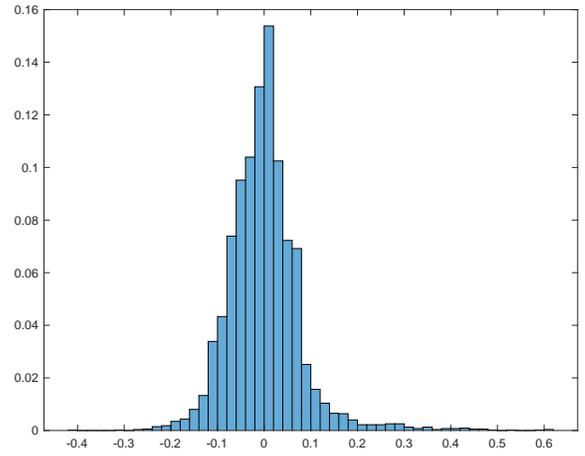
D. 1990-2001

Notes: The figure plots the distribution of expectational errors based on the Baseline information set, i.e., $\varepsilon_{i,t+1} = S_{i,t+1} - \mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^b]$. 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 $Z_{i,t} = \{ShareMfg_{i,t}, \theta_i\}$.

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 $S_{i,t}$, $S_{i,t}^{X}$, $ShareMfg_{i,t}$, and θ_i.

F   Robustness to potential model misspecifications

F.1   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 $\mathcal{I}_{i,t} = \{\theta_i, ShareMfg_{i,t}, S_{i,t}\}$.

F.1.1   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

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

Setting aside the problem that $\nu_{i,t+1}$ 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 $\rho_s > 0$, $\nu_{i,t+1}$ 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 $Z_{i,t} \nsubseteq \mathcal{I}_{i,t}$. 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.

F.1.2 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.1.1, 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.

F.1.3 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.1.1. 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.

F.2 Ideology-dependent information sets

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

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

and

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

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 $\mathcal{I}_{i,t}^b$. 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.

Figure F.1: Parameter Confidence Sets When the Rational Expectation Assumption is Violated (Expectational Errors are Correlated with $S_{i,t}$)
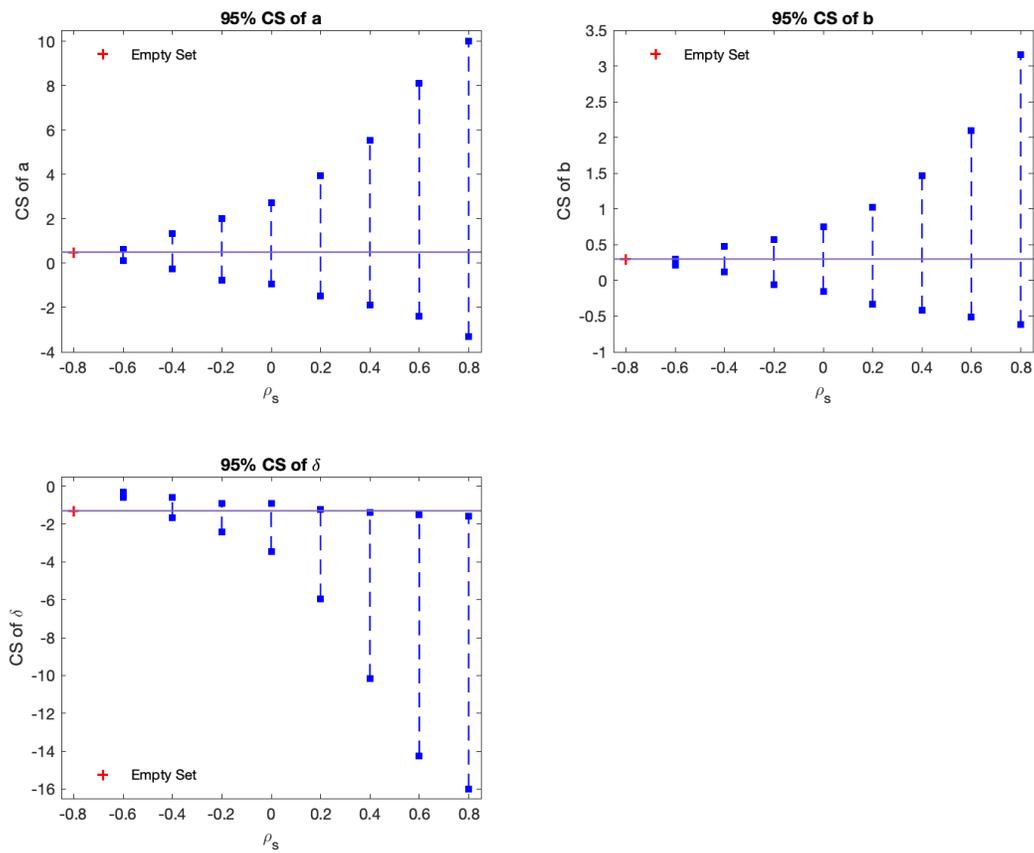
Note: The figure show the 95% confidence sets for the parameters for specifications with different ρ_s. The purple lines represent the true parameters of a, b, and δ, respectively.

Figure F.2: Parameter Confidence Sets When the Rational Expectation Assumption is Violated (Expectational Errors are Correlated with θ_i)
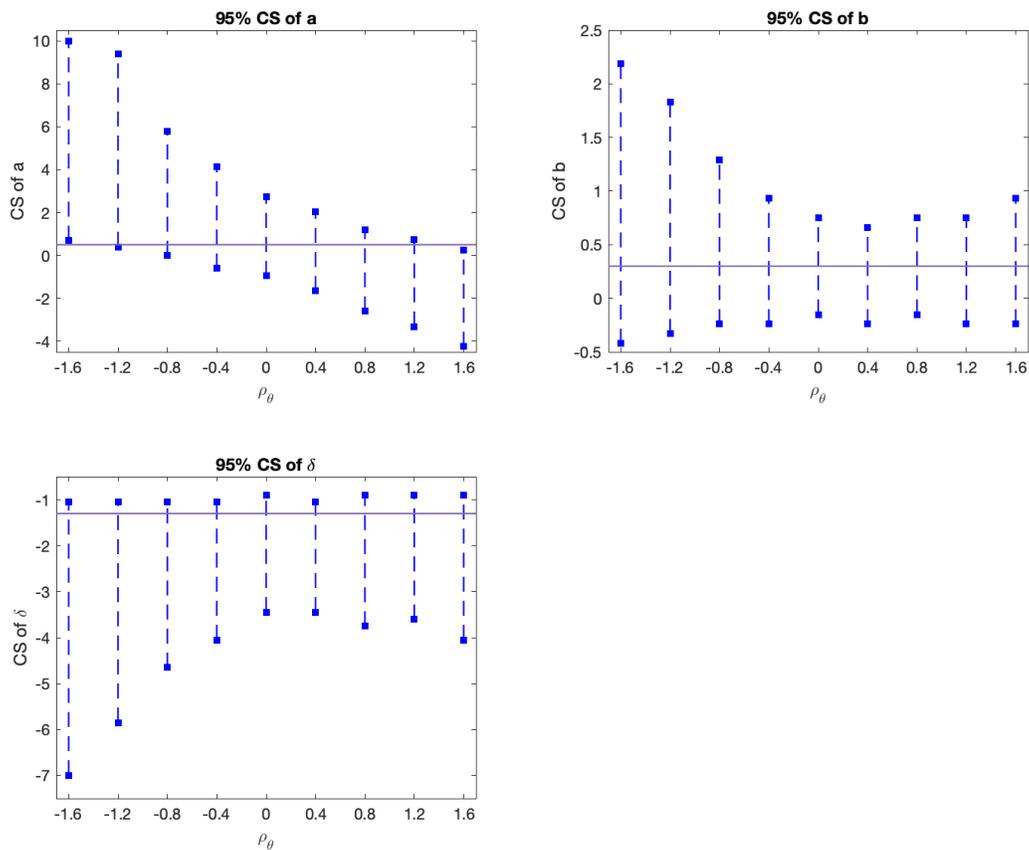
Note: The figure show the 95% confidence sets for the parameters for specifications with different ρ_θ. The purple lines represent the true parameters of a, b, and δ, respectively.

Figure F.3: Parameter Confidence Sets When the Expectation Error is a Non-Zero Constant (Expectational Errors is ρ_0)
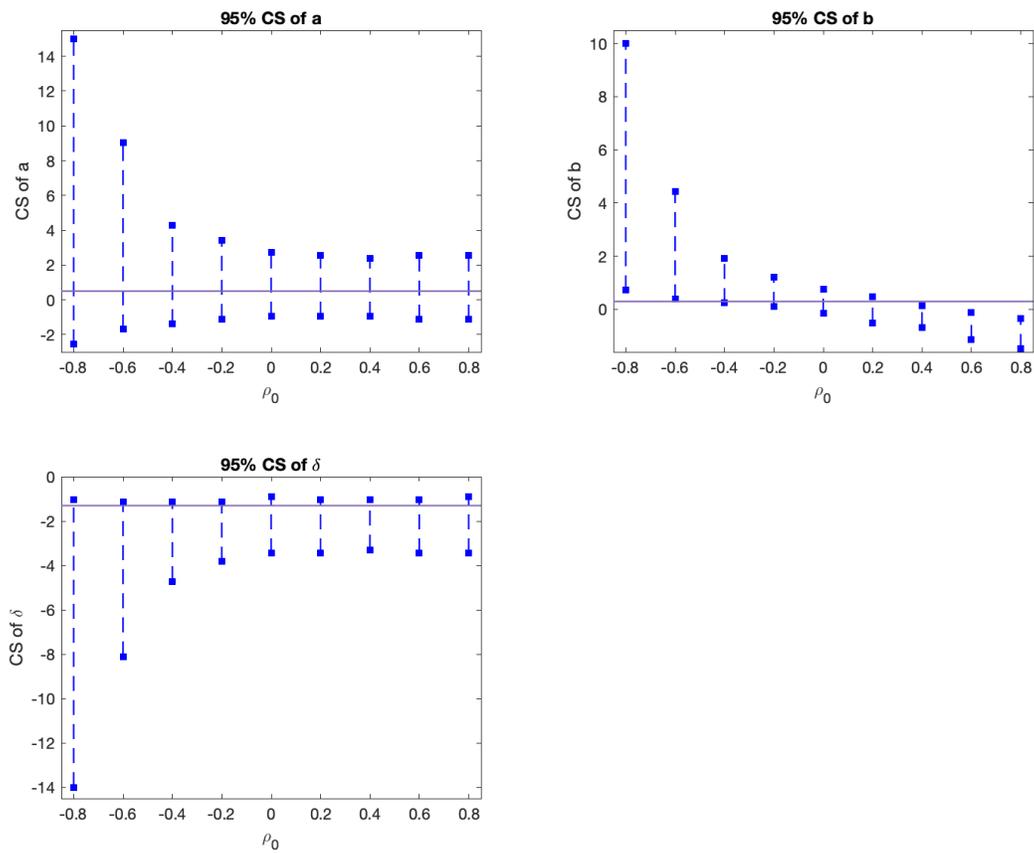
Note: The figure show the 95% confidence sets for the parameters for specifications with different ρ_θ. The purple lines represent the true parameters of a, b, and δ, respectively.

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.