Numbers We Can Believe in?
— Part 2: Putting Models to the Test —

Putting Quantitative Models to the Test: An Application to Trump's Trade War

Rodrigo Adão
Booth School of Business

Arnaud Costinot
MIT

Dave Donaldson
MIT

JIE Summer School, June 2023

Quantitative Trade Models Have Been Put to *Work*...

- **Policymakers ask: *"Should my nation change or have changed its trade policy?"***
  - Unilateral liberalization (India, Brazil...)
  - Create regional trade agreements (NAFTA, Mercosur...)
  - Amend such agreements (EU enlargement, join WTO, Brexit...)
  - Start a tariff war, retaliate against one (Trump, China...)

Quantitative Trade Models Have Been Put to *Work*...

- **Policymakers ask: *"Should my nation change or have changed its trade policy?"***
  - Unilateral liberalization (India, Brazil...)
  - Create regional trade agreements (NAFTA, Mercosur...)
  - Amend such agreements (EU enlargement, join WTO, Brexit...)
  - Start a tariff war, retaliate against one (Trump, China...)

- **Quantitative trade models provide guidance:**
  - Counterfactual simulations with and without policy change ⇒ Welfare gains or losses
  - *"After decades of supporting free trade, in 2018 the U.S. raised import tariffs and major trade partners retaliated. [...] the aggregate real income loss was $7.2 billion, or 0.04% of GDP"*

    Fajgelbaum, Goldberg, Kennedy, and Khandelwal (2020)

Quantitative Trade Models Have Been Put to *Work*...

- **Policymakers ask: *"Should my nation change or have changed its trade policy?"***
  - Unilateral liberalization (India, Brazil...)
  - Create regional trade agreements (NAFTA, Mercosur...)
  - Amend such agreements (EU enlargement, join WTO, Brexit...)
  - Start a tariff war, retaliate against one (Trump, China...)

- **Quantitative trade models provide guidance:**
  - Counterfactual simulations with and without policy change $\Rightarrow$ Welfare gains or losses
  - *"After decades of supporting free trade, in 2018 the U.S. raised import tariffs and major trade partners retaliated. [...] the aggregate real income loss was $7.2 billion, or 0.04% of GDP"*

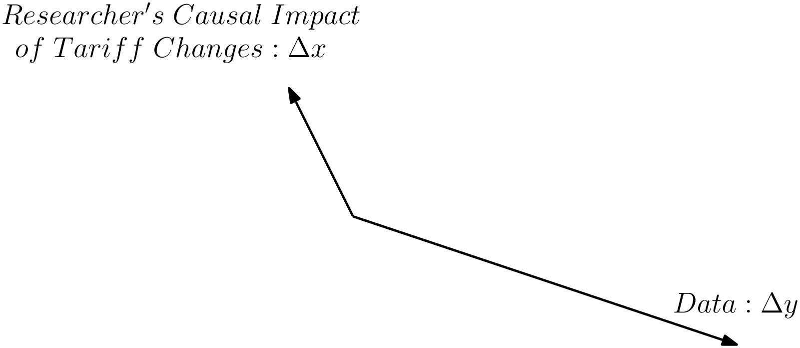
    Fajgelbaum, Goldberg, Kennedy, and Khandelwal (2020)

- **This paper's questions:**
  - How empirically credible are such counterfactual answers?
  - How can we strengthen their credibility?

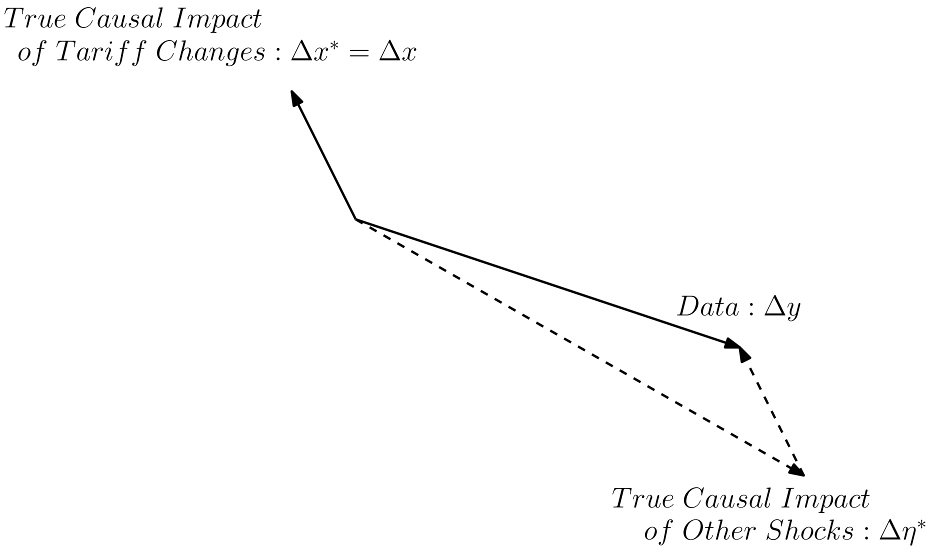
Putting Quantitative Trade Models to *Work*

$$Researcher's\ Causal\ Impact$$
$$of\ Tariff\ Changes : \Delta x$$

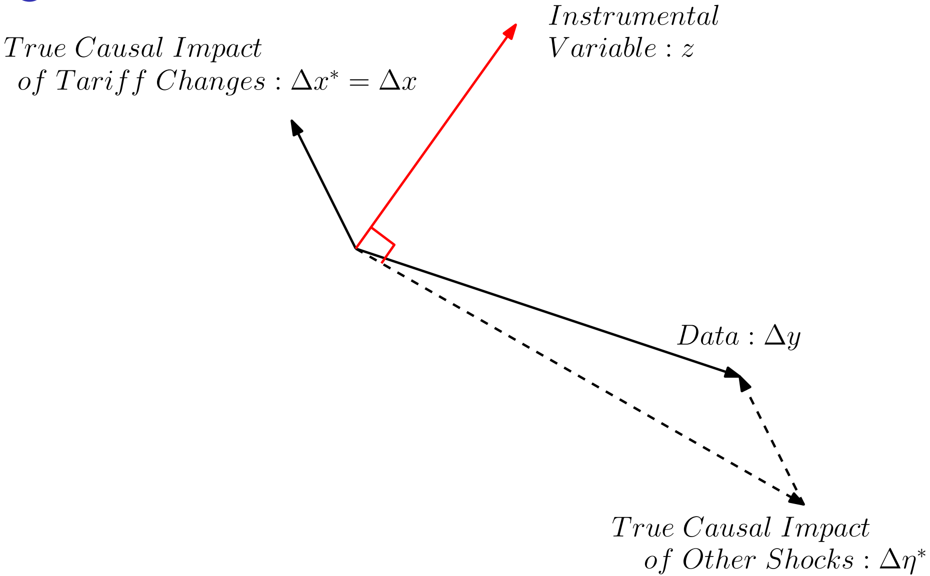
Putting Quantitative Trade Models to the *Test*



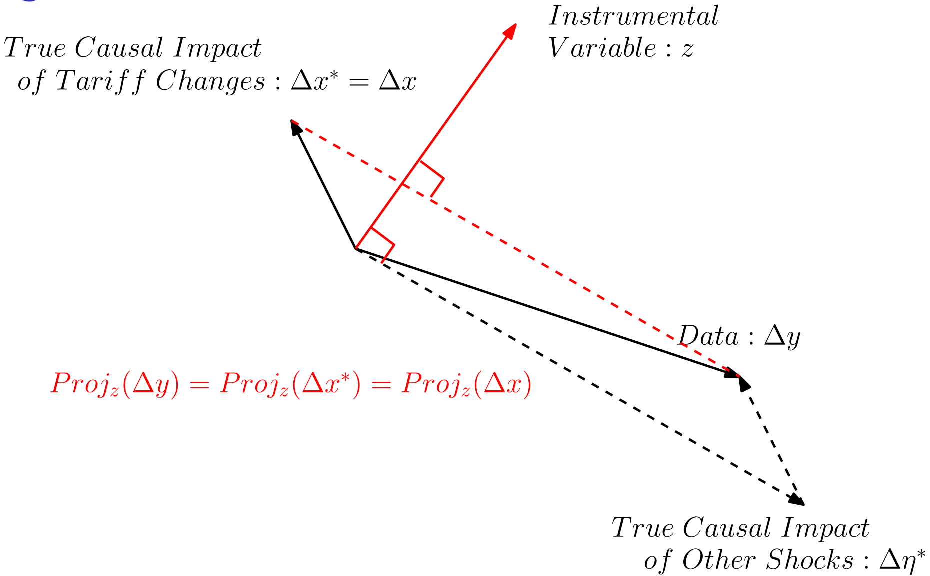
Putting Quantitative Trade Models to the *Test*



Putting Quantitative Trade Models to the *Test*



Putting Quantitative Trade Models to the *Test*



This Paper's Contribution

- **Develop an IV-based test to convey credibility of counterfactual answers:**
  1. Find exogenous policy shifters
  2. Combine exogenous policy shifters + quantitative trade model to construct IV z
  3. Compare researcher's causal impact of policy to true causal impact, up to projection on IV z
- **Advantages of IV-based test** (relative to other model validation procedures):
  1. Fully consistent with standard practices in the field:
     1.1 *Empirical side:* exogenous tariff shifters may have already been used for estimation
     1.2 *Quantitative side:* valid even when "other shocks" (not just tariffs) occur and model saturated with free parameters so as to match all data
  2. Detects and evaluates misspecification in causal impact of interest (through choice of IV)
  3. Valid despite non-trivial GE cross-sectional dependence
  4. Simple to apply $\Rightarrow$ easy *"add-on"* for existing work
- **Application: Fajgelbaum et al (2020) analysis of Trump's Trade War**
  1. Monte Carlo simulations using random shocks to assess properties of test
  2. Empirical test using actual tariff changes and observed changes in outcomes of interest

Outline

- Testing Quantitative Trade Models using IV
- Monte-Carlo Simulations
- Empirical Application to Trump's Trade War

Outline

- **Testing Quantitative Trade Models using IV**
- Monte-Carlo Simulations
- Empirical Application Trump's Trade War

Setup

- Consider reduced-form of *researcher's model*:

$$y_{n,t} = f_n(\tau_t, \epsilon_t)$$

  - $y_{n,t}$: endogenous outcome of interest n (e.g., price of a traded good)
  - τ_t: vector of all policy variables of interest (e.g., tariffs)
  - ϵ_t: vector of all time-varying parameters (e.g., technology shifters)—"other shocks"
  - $f_n(\cdot)$: mapping implied by market structure, preferences, and technology
    - both preference and technology parameters Θ may have been estimated/calibrated

- Data generated by *true model*:

$$y_{n,t} = f_n^*(\tau_t, \epsilon_t^*)$$

Credible Counterfactual Answers?

- **Causal impact of policy change:**

$$\text{Researcher's model: } \Delta x_n \equiv f_n(\tau_{t+1}, \epsilon_{t+1}) - f_n(\tau_t, \epsilon_{t+1})$$
$$\text{True Model: } \Delta x_n^* \equiv f_n^*(\tau_{t+1}, \epsilon_{t+1}^*) - f_n(\tau_t, \epsilon_{t+1}^*)$$

- **Testing as a way to enhance credibility:**

$$W(\Delta x) \equiv \sum_n \omega_n \Delta x_n \qquad \text{vs.} \qquad W(\Delta x^*) \equiv \sum_n \omega_n \Delta x_n^*$$

  - In our application, $W(\Delta x)$ is (f.o.a) of welfare change associated with Trump tariffs

- **Empirical challenge:**
  - We don't observe Δx_n^*! We observe

$$\Delta y_n = f_n^*(\tau_{t+1}, \epsilon_{t+1}^*) - f_n^*(\tau_t, \epsilon_t^*) = \Delta x_n^* + \Delta \eta_n^*$$

with $\Delta \eta_n^* \equiv f_n^*(\tau_t, \epsilon_{t+1}^*) - f_n^*(\tau_t, \epsilon_t^*)$ the causal impact of other shocks

An IV-Based Measure of Goodness of Fit

Definition

Goodness of fit of the researcher's prediction along a candidate IV z as

$$\hat{\beta}_z \equiv \frac{1}{N_W} \sum_{n \in \mathcal{N}_W} z_n (\Delta y_n - \Delta x_n),$$

where N_W denotes the number of observations in $\mathcal{N}_W \equiv \{n : \omega_n \neq 0\}$.

- **Basic Idea:** Test counterfactual answer of interest by comparing changes in data and Δy causal effect predicted by researcher's model Δx, up to projection on candidate IV z.
- **Next Step:** Show how to construct IV z such that one can characterize asymptotic distribution of $\hat{\beta}_z$ and test whether two projections coincide.

From Exogenous Policy Shifters to a Candidate IV

- Empirical literature offers exogenous policy shifters $\Delta\tau_{IV} \equiv \{\Delta\tau_{IV,k}\}$:
  - Observed tariff changes (Fajgelbaum, Goldberg, Kennedy, Khandelwal 20)
- In rest of our analysis, candidate IV z = linear function of policy shifters such that

A1 [Independence of the shifters]

Conditional on the realization of period t's tariffs and other shocks, policy shifters are independent of other shocks in period $t+1$: $\Delta\tau_{IV} \perp\!\!\!\perp \epsilon^*_{t+1} | (\epsilon^*_t, \tau_t)$.

A2 [Shift-share structure]

For any $n \in \mathcal{N}_W$, the instrumental variable takes the form $z_n = \sum_k s_{nk}\Delta\tau_{IV,k}$, where the vector of "shares" $\{s_{nk}\}$ may be a function of, and only of, the realization of period t's tariffs and other shocks, (ϵ^*_t, τ_t).

An IV-Based Test (I): Moment Restriction

An IV-Based Test (I): Moment Restriction

A3 [The causal impact of tariffs in the researcher's model is correct]

For any $n \in \mathcal{N}_W$, $\Delta x_n^* = \Delta x_n$.

An IV-Based Test (I): Moment Restriction

A3 [The causal impact of tariffs in the researcher's model is correct]

For any $n \in \mathcal{N}_W$, $\Delta x_n^* = \Delta x_n$.

Proposition 1 [Expected value of the goodness of fit measure]

Take any IV z that satisfies A1 and A2. If A3 holds, then $E_t[\hat{\beta}_z] = 0$.

An IV-Based Test (I): Moment Restriction

A3 [The causal impact of tariffs in the researcher's model is correct]

For any $n \in \mathcal{N}_W$, $\Delta x_n^* = \Delta x_n$.

Proposition 1 [Expected value of the goodness of fit measure]

Take any IV z that satisfies A1 and A2. If A3 holds, then $E_t[\hat{\beta}_z] = 0$.

- **Proof:** A1 and A2 imply $E_t[\sum_{n \in \mathcal{N}_W} z_n \Delta \eta_n^*] = 0$. Then starting from definition of $\hat{\beta}_z$ and substituting for Δy_n A3 implies

$$E_t[\hat{\beta}_z] = \frac{1}{N_W} E_t[\sum\nolimits_{n \in \mathcal{N}_W} z_n(\Delta x_n^* - \Delta x_n)] = 0$$

An IV-Based Test (I): Moment Restriction

A3 [The causal impact of tariffs in the researcher's model is correct]

For any $n \in \mathcal{N}_W$, $\Delta x_n^* = \Delta x_n$.

Proposition 1 [Expected value of the goodness of fit measure]

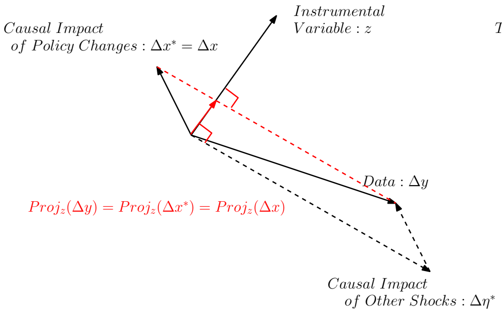
Take any IV z that satisfies A1 and A2. If A3 holds, then $E_t[\hat{\beta}_z] = 0$.

- **Proof:** A1 and A2 imply $E_t[\sum_{n \in \mathcal{N}_W} z_n \Delta \eta_n^*] = 0$. Then starting from definition of $\hat{\beta}_z$ and substituting for Δy_n A3 implies

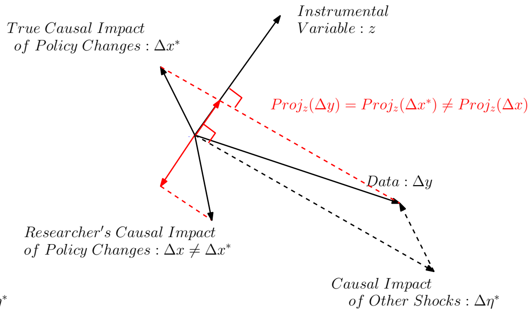
$$E_t[\hat{\beta}_z] = \frac{1}{N_W} E_t[\sum_{n \in \mathcal{N}_W} z_n (\Delta x_n^* - \Delta x_n)] = 0$$

- **Key observation:** Under A1 and A2, $E_t[\hat{\beta}_z]$ is a weighted sum of misspecifications, $\Delta x_n^* - \Delta x_n$, along all welfare-relevant variables

In a Nutshell



(a) No misspecification in causal impact



(b) Misspecification in causal impact

An IV-Based Test (II): Asymptotic Distribution

- Shift-share IV: Consistency (Borusyak et al., 2022) and Inference (Adao et al., 2019)

Proposition 2 [Asymptotic behavior of the goodness of fit measure]

Take IV z that satisfies A1 and A2. If A3 holds and (i) $\Delta\tau_{IV,k}$ are i.i.d., (ii) $\frac{1}{N_W^2}\sum_k (S_k)^2 \to 0$ with $S_k \equiv \sum_n |s_{nk}|$, and (iii) $Var_t[\Delta\tau_{IV,k}]$ and $\Delta\eta_n^*$ are uniformly bounded, then $\hat{\beta}_z \to_p 0$. If, in addition, (iv) $\frac{\max_k(S_{k,t})}{\sum_k S_{k,t}^2} \to 0$; (v) $E_t[(\Delta\tau_{IV,k})^4]$ is uniformly bounded; and (vi) $\frac{1}{\sum_k S_k^2} Var_t[\sum_{n\in\mathcal{N}_W} z_n \Delta\eta_n^* | \epsilon_{t+1}^*] \to_p V_\beta > 0$, then $r_\beta \hat{\beta}_z \to_d \mathcal{N}(0, V_\beta)$ with $r_\beta \equiv N_W/\sqrt{\sum_k S_k^2}$.

An IV-Based Test (II): Asymptotic Distribution

- Shift-share IV: Consistency (Borusyak et al., 2022) and Inference (Adao et al., 2019)

Proposition 2 [Asymptotic behavior of the goodness of fit measure]

Take IV z that satisfies A1 and A2. If A3 holds and (i) $\Delta\tau_{IV,k}$ are i.i.d., (ii) $\frac{1}{N_W^2}\sum_k (S_k)^2 \to 0$ with $S_k \equiv \sum_n |s_{nk}|$, and (iii) $Var_t[\Delta\tau_{IV,k}]$ and $\Delta\eta_n^*$ are uniformly bounded, then $\hat{\beta}_z \to_p 0$. If, in addition, (iv) $\frac{\max_k(S_{k,t})}{\sum_k S_{k,t}^2} \to 0$; (v) $E_t[(\Delta\tau_{IV,k})^4]$ is uniformly bounded; and (vi) $\frac{1}{\sum_k S_k^2} Var_t[\sum_{n\in\mathcal{N}_W} z_n \Delta\eta_n^* | \epsilon_{t+1}^*] \to_p V_\beta > 0$, then $r_\beta \hat{\beta}_z \to_d \mathcal{N}(0, V_\beta)$ with $r_\beta \equiv N_W / \sqrt{\sum_k S_k^2}$.

- **What about estimation uncertainty?** If f is known up to estimation of structural parameter θ, then we show how to compute asymptotic distribution of $\hat{\beta}_z(\hat{\theta})$ whenever
  - $\hat{\theta}$ is independent of $\hat{\beta}_z(\theta)$ (e.g. when estimation has been conducted on a different sample)
  - $\hat{\theta}$ is an IV estimator, potentially based on the same policy shifters (as in our application)

Choice of the IV (I): Economic Interpretation

- **Question:** How should we interpret goodness of fit measure? Ideally, we would like it to measure, at least on average, mispecification in the counterfactual of interest, i.e.,

$$E_t[W(\Delta x^*) - W(\Delta x)] = E_t[\sum_n \omega_n(\Delta x_n^* - \Delta x_n)]$$

Choice of the IV (I): Economic Interpretation

- **Question:** How should we interpret goodness of fit measure? Ideally, we would like it to measure, at least on average, mispecification in the counterfactual of interest, i.e.,

$$E_t[W(\Delta x^*) - W(\Delta x)] = E_t[\sum_n \omega_n(\Delta x_n^* - \Delta x_n)]$$

- **Problem:** $E_t[\hat{\beta}_z] = \frac{1}{N_W} E_t[\sum_n z_n(\Delta x_n^* - \Delta x_n)] \neq E_t[\sum_n \omega_n(\Delta x_n^* - \Delta x_n)]$ for arbitrary z
  - LATE logic with $\Delta x_n^* - \Delta x_n$ playing the role of the heterogeneous treatment

Choice of the IV (I): Economic Interpretation

- **Question:** How should we interpret goodness of fit measure? Ideally, we would like it to measure, at least on average, mispecification in the counterfactual of interest, i.e.,

$$E_t[W(\Delta x^*) - W(\Delta x)] = E_t[\sum_n \omega_n(\Delta x_n^* - \Delta x_n)]$$

- **Problem:** $E_t[\hat{\beta}_z] = \frac{1}{N_W} E_t[\sum_n z_n(\Delta x_n^* - \Delta x_n)] \neq E_t[\sum_n \omega_n(\Delta x_n^* - \Delta x_n)]$ for arbitrary z
  - LATE logic with $\Delta x_n^* - \Delta x_n$ playing the role of the heterogeneous treatment

A3'[The causal impact of tariffs in the researcher's model is misspecified]

For any $n \in \mathcal{N}_W$, $\Delta x_n^* = \alpha_n \Delta x_n$, with the misspecification parameter α_n a function of the realization of period t's tariffs and other shocks, $(\epsilon_t^*, \epsilon_t, \tau_t)$, but not period $t+1$'s tariffs, τ_{t+1}.

Choice of the IV (I): Economic Interpretation

- **Question:** How should we interpret goodness of fit measure? Ideally, we would like it to measure, at least on average, mispecification in the counterfactual of interest, i.e.,

$$E_t[W(\Delta x^*) - W(\Delta x)] = E_t[\sum_n \omega_n(\Delta x_n^* - \Delta x_n)]$$

- **Problem:** $E_t[\hat{\beta}_z] = \frac{1}{N_W} E_t[\sum_n z_n(\Delta x_n^* - \Delta x_n)] \neq E_t[\sum_n \omega_n(\Delta x_n^* - \Delta x_n)]$ for arbitrary z
  - LATE logic with $\Delta x_n^* - \Delta x_n$ playing the role of the heterogeneous treatment

A3'[The causal impact of tariffs in the researcher's model is misspecified]

For any $n \in \mathcal{N}_W$, $\Delta x_n^* = \alpha_n \Delta x_n$, with the misspecification parameter α_n a function of the realization of period t's tariffs and other shocks, $(\epsilon_t^*, \epsilon_t, \tau_t)$, but not period $t+1$'s tariffs, τ_{t+1}.

Proposition 3 [IV-based goodness of fit measures average welfare misspecification]

Take IV z that satisfies A1 and A2. If A3' holds, then one can construct z', with $z'_n \equiv z_n \omega_n E_t[\Delta x_n] / E_t[z_n \Delta x_n]$ for all $n \in \mathcal{N}_W$, such that $E_t[\hat{\beta}_{z'}] = E_t[W(\Delta x^*) - W(\Delta x)]$.

Choice of the IV (II): Statistical Power

- Three potential reasons for low-power of arbitrary IV-based test:
  1. **Lack of first stage:** $E_t[z_n \Delta x_n] = E_t[z_n \Delta y_n] = 0$ because z is noise
  2. **Mechanical fit:** Estimation moments mechanically related to testing moments
  3. **Precision:** Too much variance in $\Delta y_n - \Delta x_n \Rightarrow$ too much variance in $\hat{\beta}_z$

Choice of the IV (II): Statistical Power

- Three potential reasons for low-power of arbitrary IV-based test:
    1. **Lack of first stage:** $E_t[z_n \Delta x_n] = E_t[z_n \Delta y_n] = 0$ because z is noise
    2. **Mechanical fit:** Estimation moments mechanically related to testing moments
    3. **Precision:** Too much variance in $\Delta y_n - \Delta x_n \Rightarrow$ too much variance in $\hat{\beta}_z$
- Three potential solutions:
    1. To address lack of first stage, **use** causal impact of shifters predicted by **researcher's model,** i.e. $s_{nk} = \partial f_n / \partial \tau_k \Rightarrow z_n = \sum_k (\partial f_n / \partial \tau_k) \Delta \tau_{IV,k}$
    2. To address mechanical fit, use IV z such that **estimation moments** are **less informative** about $\hat{\beta}_z$ in the sense of Andrews et al. (2020)
    3. To address precision, project z on a vector of controls and **use residuals**

Other Tests: A Moment is a Moment?

- We have proposed the test statistic $\hat{\beta}_z$ as a way to test causal mechanisms
  - It asks:*"can we reject the null that* $W(\Delta x) = W(\Delta x^*)$*?"*
  - It does not ask: *"can we reject the null that* $f = f^*$*?"*
- A number of popular approaches to model validation do not draw this distinction:

Other Tests: A Moment is a Moment?

- We have proposed the test statistic $\hat{\beta}_z$ as a way to test causal mechanisms
  - It asks:*"can we reject the null that $W(\Delta x) = W(\Delta x^*)$?"*
  - It does not ask: *"can we reject the null that $f = f^*$?"*
- A number of popular approaches to model validation do not draw this distinction:

1. **Correlation-based test** – e.g., Kehoe (05); Kehoe, Pujolas, Rossbach (17)
   - Use Δy_n as proxy for Δx_n^* and compute $corr(\Delta y_n, \Delta x_n)$ (or R^2 of reg of Δy_n on Δx_n)
   - **Problem:** Evaluates importance of other shocks for Δy_n, not impact of τ_t on y_t since $\text{corr}(\Delta y_n, \Delta x_n) \propto \frac{\text{var}(\Delta x_n^*)}{\text{var}(\Delta \eta_n^*)}$
2. **Untargeted moments from the initial equilibrium** – e.g., Edmond, Midrigan, Xu (11); Costinot, Donaldson, Smith (16); Antras Fort Tintelnot (17)
   - Pick parameters Θ to match a set $\{y_{n,t}\}_{n\in\mathcal{N}_T}$ of targeted equilibrium outcomes at t
     - Report a good fit for a set $\{y_{n,t}\}_{n\in\mathcal{N}_U}$ of untargeted equilibrium outcomes at t
     - **Problem:** Evaluates relationship between $\{y_n\}_{n\in\mathcal{N}_T}$ and $\{y_n\}_{n\in\mathcal{N}_U}$, not impact of τ_t on y_t

Other Tests: A Causal Moment is a Causal Moment?

- **Lucas (1980) program:**

  "need to test them (models) as useful imitations of reality by subjecting them to shocks for which we are fairly certain how actual economies or parts of economies would react. The more dimensions on which the model mimics the answers actual economies give to simple questions, the more we trust its answers to harder questions."

Other Tests: A Causal Moment is a Causal Moment?

- **Lucas (1980) program:**

  "need to test them (models) as useful imitations of reality by subjecting them to shocks for which we are fairly certain how actual economies or parts of economies would react. The more dimensions on which the model mimics the answers actual economies give to simple questions, the more we trust its answers to harder questions."

- **Examples:** monetary shocks (Christiano et al., 2005), government spending (Nakamura and Steinsson 2014), Berlin wall (Ahlfeldt et al. ,2015)
- IV-based test part of same broad program, but not all "dimensions on which the model mimics actual economies" are made equal
  - Which outcome variable Δy_n should one focus on? Which IV z_n should one use to improve power and avoid mechanical fit? How should one conduct inference in the presence of GE linkages and prior estimation? And how should the test statistic be interpreted?

Other Tests: A Causal Moment is a Causal Moment?

- **Lucas (1980) program:**

  "need to test them (models) as useful imitations of reality by subjecting them to shocks for which we are fairly certain how actual economies or parts of economies would react. The more dimensions on which the model mimics the answers actual economies give to simple questions, the more we trust its answers to harder questions."

- **Examples:** monetary shocks (Christiano et al., 2005), government spending (Nakamura and Steinsson 2014), Berlin wall (Ahlfeldt et al. ,2015)
- IV-based test part of same broad program, but not all "dimensions on which the model mimics actual economies" are made equal
  - Which outcome variable Δy_n should one focus on? Which IV z_n should one use to improve power and avoid mechanical fit? How should one conduct inference in the presence of GE linkages and prior estimation? And how should the test statistic be interpreted?
  - We offer answers to all four questions

Outline

- Testing Quantitative Trade Models using IV
- **Monte-Carlo Simulations**
- Empirical Application to Trump's Trade War

The Researcher's Model: FGKK (2020)

- Home ($i = H$) and its trading partners ($i \in \mathcal{I}$)
- Multiple sectors, $s \in \mathcal{S}$
- Time is discrete and indexed by t

Domestic Households

- Household in region r and sector s has endowment of $L_{rs,t}$ units of labor
- All households have common nested CES preferences:

$$U_t = (C_{NT,t})^{\beta_{NT,t}}(C_{T,t})^{\beta_{T,t}}$$

$$C_{T,t} = \prod_{s \in \mathcal{S}} (C_{Ts,t})^{\beta_{s,t}}, \qquad C_{Ts,t} = \left[(A_{Ds,t})^{\frac{1}{\kappa}} (D_{s,t})^{\frac{\kappa-1}{\kappa}} + (A_{Ms,t})^{\frac{1}{\kappa}} (M_{s,t})^{\frac{\kappa-1}{\kappa}} \right]^{\frac{\kappa}{\kappa-1}}$$

$$D_{s,t} = \left[\sum_{g \in \mathcal{G}_s} (a_{Dg,t})^{\frac{1}{\eta}} (d_{g,t})^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}, \qquad M_{s,t} = \left[\sum_{g \in \mathcal{G}_s} (a_{Mg,t})^{\frac{1}{\eta}} (m_{g,t})^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}$$

$$m_{g,t} = \left[\sum_i (a_{ig,t})^{\frac{1}{\sigma}} (m_{ig,t})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

Domestic Firms

- Competitive firms in each region r and sector s take good and factor prices as given
- Nested CES technologies:

$$Q_{NTr,t} = Z_{NTr,t} L_{NTr,t}$$

$$Q_{sr,t} = Z_{sr,t} \left(I_{sr,t}\right)^{\alpha_{Is,t}} \left(L_{sr,t}\right)^{\alpha_{Ls,t}}, \qquad \alpha_{Is,t} + \alpha_{Ls,t} < 1$$

$$I_{sr,t} = \prod_{k \in \mathcal{S}} (I_{ksr,t})^{\alpha_{ks,t}}, \qquad \sum_{k \in \mathcal{S}} \alpha_{ks,t} = 1$$

$$\sum_{g \in \mathcal{G}_s} \frac{q_{gs,t}}{z_{gs,t}} = \sum_{r} Q_{sr,t}$$

Foreign Import Demand and Export Supply

- Given export price $p^x_{ik,t}$, exports (given by foreign import demand):

$$x_{ig,t} = a^*_{ig,t}\left((1+\tau^*_{ig,t})p^X_{ig,t}\right)^{-\sigma^*}$$

Foreign Import Demand and Export Supply

- Given export price $p^{x}_{ik,t}$, exports (given by foreign import demand):

$$x_{ig,t} = a^{*}_{ig,t}\left((1+\tau^{*}_{ig,t})p^{X}_{ig,t}\right)^{-\sigma^{*}}$$

- Given (pre-tariff) import price $p^{*}_{ig,t}$, foreign export supply is

$$m_{ig,t} = (p^{*}_{ig,t})^{\frac{1}{\omega^{*}}}(z^{*}_{ig,t})^{\frac{1}{\omega^{*}}}$$

Foreign Import Demand and Export Supply

- Given export price $p^x_{ik,t}$, exports (given by foreign import demand):

$$x_{ig,t} = a^*_{ig,t}\left((1+\tau^*_{ig,t})p^X_{ig,t}\right)^{-\sigma^*}$$

- Given (pre-tariff) import price $p^*_{ig,t}$, foreign export supply is

$$m_{ig,t} = (p^*_{ig,t})^{\frac{1}{\omega^*}}(z^*_{ig,t})^{\frac{1}{\omega^*}}$$

- Government at Home imposes import tariffs so that import price is

$$p_{ig,t} = (1+\tau_{ig,t})p^*_{ig,t}$$

Foreign Import Demand and Export Supply

- Given export price $p^{x}_{ik,t}$, exports (given by foreign import demand):

$$x_{ig,t} = a^{*}_{ig,t}\left((1+\tau^{*}_{ig,t})p^{X}_{ig,t}\right)^{-\sigma^{*}}$$

- Given (pre-tariff) import price $p^{*}_{ig,t}$, foreign export supply is

$$m_{ig,t} = (p^{*}_{ig,t})^{\frac{1}{\omega^{*}}}(z^{*}_{ig,t})^{\frac{1}{\omega^{*}}}$$

- Government at Home imposes import tariffs so that import price is

$$p_{ig,t} = (1+\tau_{ig,t})p^{*}_{ig,t}$$

- Government uses a lump-sum transfer T_t to rebate tariff revenue and foreign transfer D_t

Back to General Notation

- Time-varying shocks to preferences, technology, and endowments:

$$\epsilon_t \equiv \{\beta_{NT,t}, \beta_{s,t}, A_{Ms,t}, a_{Dg,t}, a_{Mg,t}, a_{ig,t}, Z_{NTr,t}, Z_{sr,t}, \alpha_{Is,t}, \alpha_{Ls,t}, \alpha_{ksr,t}, a^*_{ig,t}, z^*_{ig,t}, D_t, L_{sr,t}\}$$

- Governments' policy vector:

$$\tau_t \equiv \{\tau_{ig,t}, \tau^*_{ig,t}\}$$

- UMP + PMP + GMC + LMC + GBC $\Longrightarrow$ **reduced-form** $y_t = f(\tau_t, \epsilon_t)$

Causal Effect of Interest

- Welfare impact of tariff changes (up to first-order):

$$W(\Delta x) = \sum_{i,v} [\omega_{iv}^X (\Delta x_{iv}^X) - \omega_{iv}^M (\Delta x_{iv}^M) + \omega_{iv}^T (\Delta x_{iv}^T)],$$

  where:
  - $\Delta x_{iv}^X \equiv$ changes in the log of US export prices of good v in country i (pre-foreign tariff)
  - $\Delta x_{iv}^M \equiv$ changes in the log of US import prices of good v from country i (post-US tariff)
  - $\Delta x_{iv}^T \equiv$ changes in US tariff revenues on good v from country i (as share of import spending)
  - $\omega_{iv}^X \equiv$ share of export revenues in 2016 US GDP accounted by country i and good v
  - $\omega_{iv}^M = \omega_{iv}^T \equiv$ share of import spending in 2016 US GDP accounted by country i and good v

Calibration of FGKK's Model

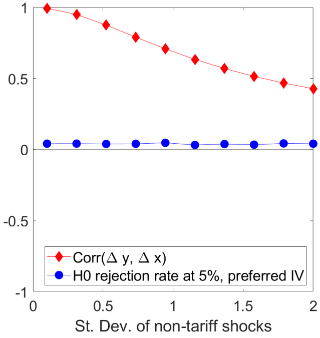
- Simulations to match setting studied by FGKK
  - Foreign countries i: 71 partner countries (99% of trade)
  - Sectors s: 88 tradable (NAICS); Products g: 10,228 tradable (10-digit HS)
- Elasticities, Θ, given by FGKK's parameter estimates:
  - $\kappa = 1.19, \eta = 1.53, \sigma = 2.53, \sigma^* = 1.04, \omega^* = -0.002$
- Given Θ and τ_t, back out $\epsilon_t = f^{-1}(\tau_t, y_t|\Theta)$ from FGKK's data on:
  - variety-level exports/imports
  - sector-level output, labor, intermediate spending, final sales
  - county-sector employment

Simulation Procedure

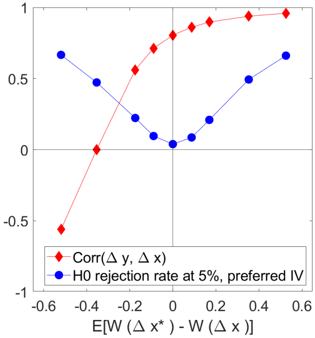
- We simulate the economic impact of import tariffs and other shocks
- For each simulation step $b = 1...B$ (with $B = 2500$):
  1. Compute counterfactual prices and quantities y_{t+1}^b from true model—either FGKK's model or a misspecified version of this model—given
     - Independent random draws of changes in tariffs $\Delta\tau^b$ from a log-normal distribution
     - Independent random draws of $\{\Delta a_{ig}^*, \Delta z_{ig}^*, \Delta a_{ig}\}^b$ from a log-normal distribution
  2. Given counterfactual prices and quantities y_{t+1}^b
     - Compute welfare impact of tariffs in true model $[W(\Delta x^*)]^b$ + researcher's model $[W(\Delta x)]^b$
     - Compute our goodness-of-fit measure $(\hat{\beta}_z)^b$ for various IVs z
     - Also compute correlation between $(\Delta y)^b$ and $(\Delta x)^b$.

Correlation- versus IV-based Tests

"Preferred IV" starts from $s_{nk} = \partial f_n / \partial \tau_k$ and applies Proposition 3's adjustment



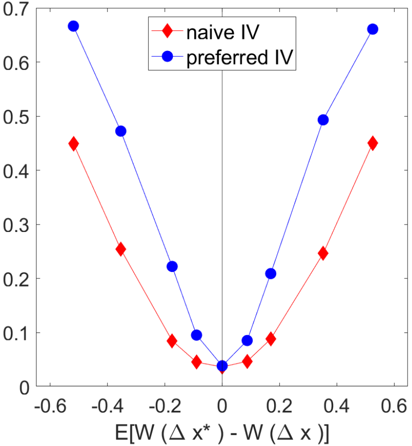
(c) No Misspecification



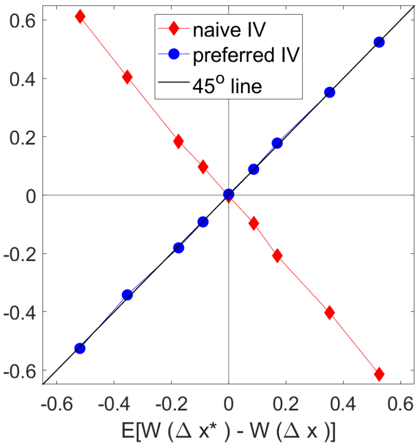
(d) Uniform Misspecification

Comparing IV-Based Tests (Uniform Misspecification)

"Preferred IV" as before. "Naive IV" only uses tariff shifters on product of interest.



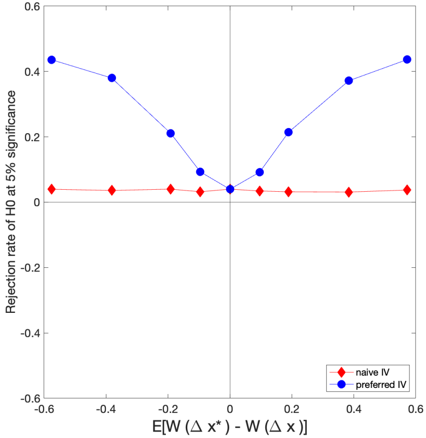
(e) Rejection rate



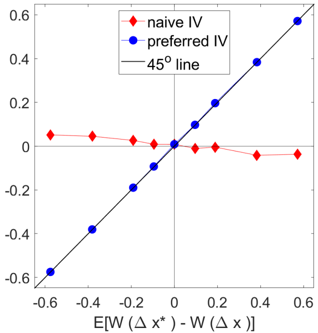
(f) Mean goodness-of-fit value

Comparing IV-Based Tests (Export Misspecification)

"Preferred IV" and "Naive IV" as before.



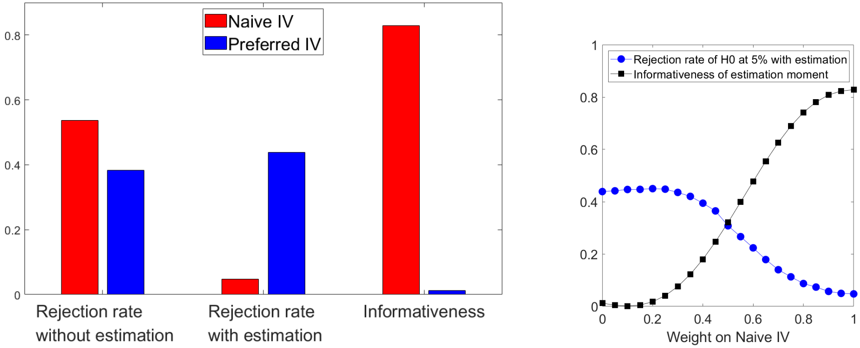
(g) Rejection rate



(h) Mean goodness-of-fit value

Estimation, Informativeness and Mechanical Fit

"Preferred IV" as before. "Naive IV" further residualized with respect to product-specific fixed effects. σ estimated as in FGKK using product-specific fixed effects. Import quantities are misspecified.



Outline

- Testing Quantitative Trade Models using IV
- Monte-Carlo Simulations
- **Empirical Application to Trump's Trade War**

Empirical Application: Trump's Trade War, 2016-2019

- Everything exactly as in previous simulations, except:

  1. Use actual US and foreign tariff changes:
     - $\tau_t \equiv \{\tau_{ig,t}, \tau^*_{ig,t}\}$: avg. Jan-Dec, 2016
     - $\tau_{t+1} \equiv \{\tau_{ig,t+1}, \tau^*_{ig,t+1}\}$: avg. Jan-April, 2019

  2. Use actual data on post-shock outcomes y_{t+1}

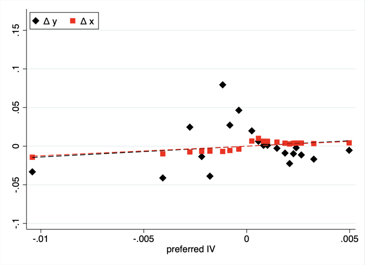
     - Given Θ and τ_{t+1}, back out $\epsilon_{t+1} = f^{-1}(\tau_{t+1}, y_{t+1}|\Theta)$ from FGKK's data

Putting FGKK to the (IV-Based) test

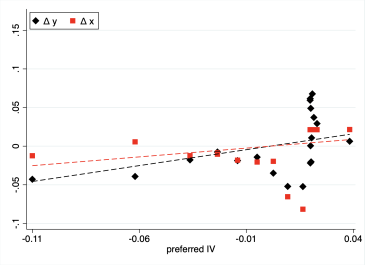
Goodness of fit measure:	Correlation	IV-Based Test	
		Naive IV	Preferred IV
	$Corr\left(\Delta y_n, \Delta x_n(\hat{\theta})\right)$	$\hat{\beta}_{z^{\text{naive}}}(\hat{\theta})$	$\hat{\beta}_{z^{\text{pref}}}(\hat{\theta})$
	(1)	(2)	(3)
Point estimate	0.08	-0.01	-0.09
Inference ignoring estimation uncertainty			
Std. error		0.18	0.15
p-value of H0: $\hat{\beta} = 0$		0.96	0.56
Inference accounting for estimation uncertainty			
Std. error		0.24	0.18
p-value of H0: $\hat{\beta} = 0$		0.97	0.63

A Final Diagnosis

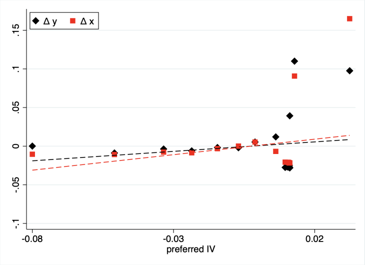
Observed and Predicted changes against “Preferred IV”



(k) Export Prices



(l) Import Prices



(m) Tariff Revenues

Concluding Remarks

- **Goal of paper:** develop an IV-based test that can help users of quantitative models to convey credibility of their counterfactual answers of interest
- **Key requirement:** exogenous policy shifters (but can be those used for model estimation)
- **Benefits:**
  - Works even after structural estimation
  - Works in presence of "other shocks" of arbitrary sizes
  - Works when model calibrated so as to exactly fit data
  - Valid inference despite complicated GE dependence across observations tested
  - Can tailor test to focus on counterfactual question of interest
  - Easy to apply: only extra requirement is to compute model's Jacobian