Quotas in General Equilibrium

David Baqaee and Kunal Sangani

May 2025

Quota Distortions

- Standard approach models misallocation using implicit taxes or "wedges."
- Wedges natural in some contexts (ad valorem taxes, markups, sticky prices).
- In other contexts, distortions directly constrain quantities without regard to prices.
  - E.g., import quotas, visa caps, zoning restrictions, emissions limits, local content requirements, land use ceilings, taxicab medallions.
  - Missing markets (land markets, credit markets, insurance markets).
- This paper: A general framework for analyzing economies with quota-like distortions.

Preview of Results

- Result: Any feasible allocation can be implemented as the decentralized equilibrium of an economy with quotas.
- Implication 1: Can study any distorted allocation of resources using (implicit) quotas.
- Implication 2: Since resulting equilibrium is constrained efficient, greatly simplifies comparative statics.
  - Not subject to Theory of Second Best.
  - Comparative statics disciplined by small set of sufficient statistics.
  - Can further characterize distance to efficient frontier, nonlinearities away from the frontier.

Preview of Results

- Response of output to quota changes and productivity shocks.
  - Key statistic: Profits earned by quota holders.
  - Don’t need elasticities of substitution / production, full input-output structure, etc.
  - Examples: H-1B visa cap, zoning restrictions on single-family housing.

Preview of Results

- Response of output to quota changes and productivity shocks.
  - Key statistic: Profits earned by quota holders.
  - Don't need elasticities of substitution / production, full input-output structure, etc.
  - Examples: H-1B visa cap, zoning restrictions on single-family housing.
- Nonlinearities, distance to the frontier.
  - Concavity/convexity depends on how profits change in response to shocks.
  - Nonlinearities can be expressed in terms of quota demand system.
  - Distance to frontier depends on (1) profits and (2) elasticity of profits to quota.
  - Examples: Argentina capital controls, textile & clothing quotas, NYC taxicab medallions.

Selected Related Literature

- **Misallocation with wedges.**
  - E.g., Harberger (1954), Basu and Fernald (2002), Chari et al. (2007), Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Petrin and Levinsohn (2012), Baqaee and Farhi (2020), Bigio and La'O (2020), Edmond et al. (2023).
- **Microeconomic shocks and aggregate efficiency.**
  - Efficient economies: Domar (1961), Hulten (1978), Foerster et al. (2011), Gabaix (2011), Acemoglu et al. (2012), Atalay (2017), Baqaee and Farhi (2019).
  - Inefficient economies: Baqaee (2018), Grassi (2017), Liu (2019), Reischer (2019), Baqaee and Farhi (2020), Buera and Trachter (2024).
- **Studies of specific quantity-based distortions.**
  - (*Trade*): Feenstra (1988), Feenstra (1992), Brambilla et al. (2010), Khandelwal et al. (2013), (*Housing*): Glaeser and Gyourko (2018), Hsieh and Moretti (2019).
- Less related: Public finance literature on using quotas vs. taxes to achieve policy objectives. (E.g., Weitzman 1974, Dasgupta and Stiglitz 1977).

Table of Contents

Table of Contents

General Framework

- Representative household, N goods indexed by i, F factors indexed by f.
- Real output Y maximizes constant-returns aggregator $\mathscr{D}$,

$$Y = \max_{\{c_1,\ldots,c_N\}} \mathscr{D}(c_1, \ldots, c_N),$$

subject to the budget constraint,

$$\sum_{i}^{N} p_i c_i = \sum_{f=1}^{F} w_f L_f + \sum_{i=1}^{N} \Pi_i,$$

where c_i final demand, p_i prices, L_f factor endowments, w_f wages, and Π_i profits.

General Framework: Quotas

- Each good i produced using constant returns production technology,

$$A_i F_i(x_{i1}, \ldots, x_{iN}, L_{i1}, \ldots, L_{iF}),$$

where x_{ij} is use of intermediate good j, L_{if} use of factor f, and A_i productivity shifter.

- A quota restricts the production of good i at a quantity y_{i^*},

$$y_i = \min\{y_{i^*}, A_i F_i(x_{i1}, \ldots, x_{iN}, L_{i1}, \ldots, L_{iF})\}.$$

- Profits for producers of i are revenues less intermediate and factor costs,

$$\Pi_i = p_i y_i - \sum_{i=1}^{N} p_j x_{ij} - \sum_{f=1}^{F} w_f L_{if}.$$

- Take nominal GDP as numeraire, $\sum_i p_i c_i = 1$. Denote Domar weights $\lambda_i \equiv p_i y_i$.

Equilibrium

- Given quotas y_{i^*}, productivities A_i, production functions F_i, and factor endowments L_f, an equilibrium consists of prices p_i, wages w_f, outputs y_i, final demands c_i, and intermediate / factor input choices x_{ij} and L_{if} such that:
  - Final demands c_i maximizes real output subject to the budget constraint.
  - Each producer maximizes profits, taking prices and quotas as given.
  - For all goods with quotas, $y_i \leq y_{i^*}$.
  - Resource constraints satisfied:

$$c_i + \sum_{j=1}^{N} x_{ji} \leq y_i \text{ for all } i \qquad \text{and} \qquad \sum_{i=1}^{N} L_{if} \leq L_f \text{ for all } f.$$

Implementing an Allocation Using Quotas

Definition (Feasible allocation)

An allocation $(\{c_i\}, \{x_{ij}\}, \{L_{if}\})$ is feasible if:

- c_i, x_{ij}, and L_{if} are non-negative for all i, j, and f,
- $y_i \leq A_i F_i(x_{i1}, \ldots, x_{iN}, L_{i1}, \ldots L_{iF})$ for all i,
- Resource constraints are satisfied.

Proposition

Consider some feasible allocation $\mathscr{X}$. Then:

1. *There exists a vector of quotas, $\{y_{i^*}\}$, such that the equilibrium has allocation $\mathscr{X}$.*
2. *Given these quotas, the allocation $\mathscr{X}$ is efficient.*

Implementing an Allocation Using Quotas

Proposition

Consider some feasible allocation $\mathscr{X}$. Then:

1. *There exists a vector of quotas, $\{y_{i^*}\}$, such that the equilibrium has allocation $\mathscr{X}$.*
2. *Given these quotas, the allocation $\mathscr{X}$ is efficient.*

- Add nodes/quotas to guarantee that competitive eqm. yields desired allocation.
- First Welfare Theorem implies allocation is constrained efficient.
- Implication 1: Recast any distorted allocation as eqm. of economy with implicit quotas.
- Implication 2: Analyze eqm. using tools for efficient economies (e.g., Hulten's Thm).

Table of Contents

Table of Contents

First-Order Effects

Proposition (First-Order Effects in Economies with Quotas)

To a first order, the effect of changes in quotas y_{i^*} and productivities A_i on output is

$$d\log Y = \sum_i \Pi_i d\log y_{i^*} + \sum_i (\lambda_i - \Pi_i)\, d\log A_i,$$

where λ_i and Π_i are sales and profits of i divided by GDP.

If all quotas are non-binding, then $d\log Y = \sum_i \lambda_i d\log A_i$.

- Profits of constrained producers are sufficient statistic for effect of quota changes.
- Holding other quotas fixed, removing quota always improves welfare (no Second Best).
  - If quotas adjust endogenously, Theory of Second Best returns.
- When all quotas are non-binding, profits are zero $\Rightarrow$ Hulten's Theorem.

Comparison to Wedge Distortions

- Equilibrium with wedges: Given wedges τ_i, the price of good i equals its marginal cost times τ_i; wedge revenues $\Pi_i = (1 - 1/\tau_i)\, p_i y_i$ rebated to representative household, ...

Proposition (First-Order Effects in Economies with Wedges)

To a first order, the effect of changes in wedges τ_i and productivities A_i on output is

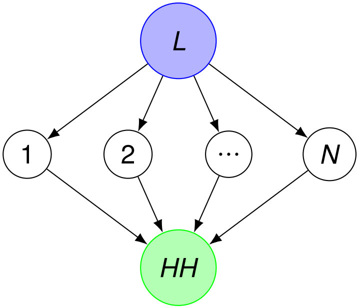
$$d\log Y = \sum_i \sum_j \Pi_i \left[\frac{\partial \log y_i}{\partial \log \tau_j} d\log \tau_j + \frac{\partial \log y_i}{\partial \log A_j} d\log A_j \right] + \sum_i (\lambda_i - \Pi_i)\, d\log A_i,$$

where $\partial \log y_i / \partial \log \tau_j$ and $\partial \log y_i / \partial \log A_j$ are general-equilibrium elasticities of y_i with respect to changes in τ_j and A_j respectively.

- Without initial distortions, results coincide with quota economies / Hulten's theorem.
- Otherwise, depends on how all producers' quantities adjust $\Rightarrow$ Theory of Second Best.

Illustrative Example 1: Reallocations Under Quotas vs. Wedges

Figure: Horizontal economy.



- Quotas: effect of relaxing quota on firm 1:

$$d\log Y = \Pi_1 d\log y_{1^*}.$$

Resources always come from unconstrained users.

- Wedges: effect of reducing τ_1 to increase output by $d\log y_1$ is

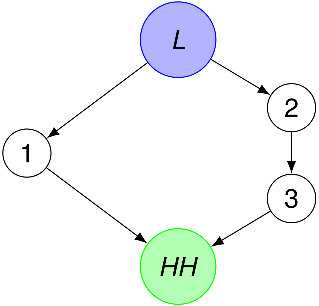
$$d\log Y = \Pi_1 d\log y_1 - \frac{L_1}{L - L_1}(\bar{\Pi} - \Pi_1) d\log y_1.$$

where $\bar{\Pi}$ is the economy-wide profit share.

- Resources reallocated from all other firms, including more distorted users. Need to know Π_1 relative to $\bar{\Pi}$.

Illustrative Example 2: Interdependent Producers

Figure: Interdependent firms.



- What are the gains from relaxing distortion on firm 2?
- Wedges:

$$d \log Y = \sum_i \Pi_i \frac{\partial \log y_i}{\partial \log \tau_2} d \log \tau_2$$
$$= \theta \left[\Pi_1 - (\Pi_1 + \Pi_2 + \Pi_3)(L_1/L)\right] d \log \tau_2 .$$

- Reducing τ_2 not always beneficial, and comparing $\Pi_2 \lesseqgtr \Pi_1$ not sufficient to know whether beneficial.
- Quotas: Reducing distortion always increases output.

$$d \log Y = \Pi_2 d \log y_{2^*} .$$

Table of Contents

Empirical Example 1: H-1B Visa Quota

- H-1B visa for high-skill foreign workers capped at 85,000 visas/year since mid-2000s.
  - Applications each year exceed cap, so visas awarded through a lottery.

- To a first order, gains from increasing quota equal to rents earned by visa winners:

$$d \log Y = \Pi_i d \log y_{i^*} \approx \frac{\Pi_i}{y_{i^*}} dy_{i^*}.$$

- Clemens (2013) compares earnings of winners vs. losers of 2007 H-1B lottery.
  - Earnings for workers who won lottery were \$12,641 higher two years after the lottery.
  - Assuming workers paid their marginal product, difference in earnings isolates excess profits of workers with right to migrate.

- Doubling number of visas in 2007 would have increased world output \$1.07B.

Empirical Example 2: Zoning Restrictions on Single-Family Housing

- What are the gains from loosening zoning restrictions on single-family housing?

- To a first order, given by profits of rights to build new single-family housing.
  - Gyourko and Krimmel (2021) isolate “zoning taxes” by comparing land value for parcels with rights to build new single-family housing to value of land with existing housing.

- Note: Efficiency gains expressed directly in terms of new units permitted.
  - Wedge approach would require mapping quantities into changes in effective zoning tax.

- Applies even in economies with other (quantity-based) distortions.

Empirical Example 2: Zoning Restrictions on Single-Family Housing

- What are the gains from loosening zoning restrictions on single-family housing?

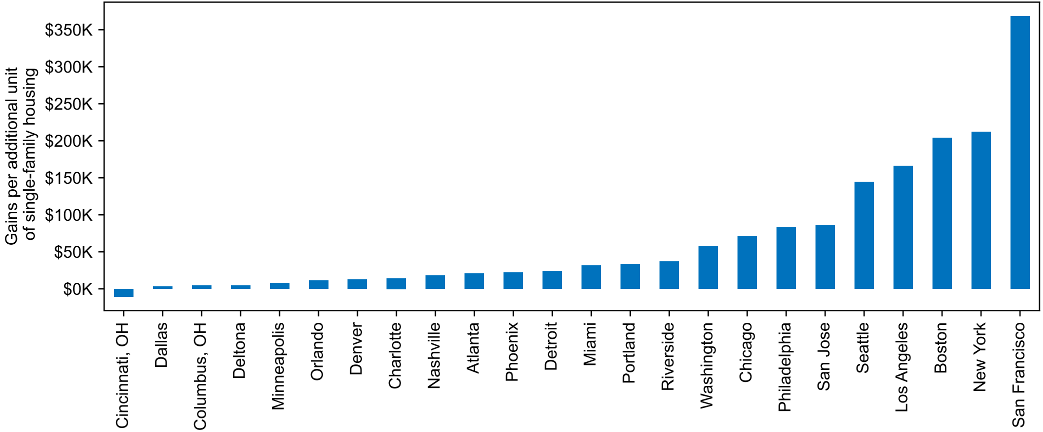


Table of Contents

Table of Contents

Nonlinearities

Proposition (Nonlinear Effects of Single Quota)

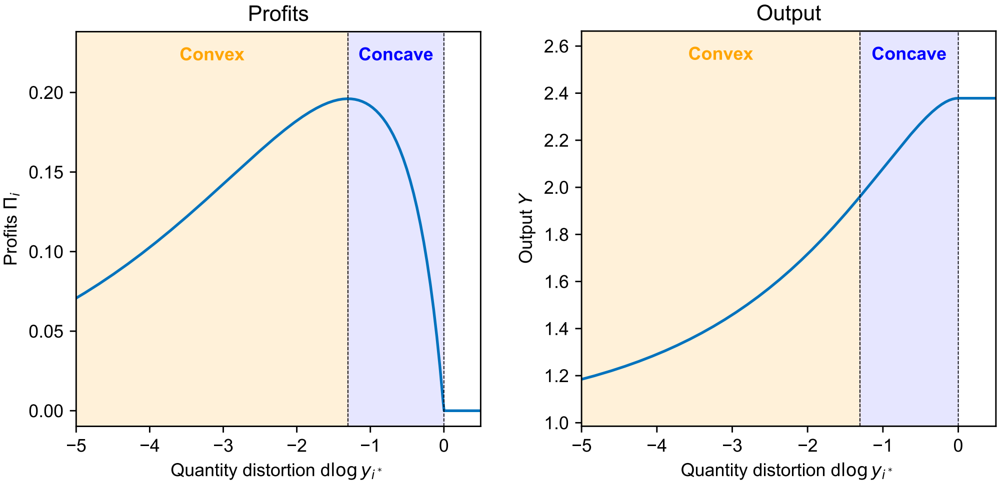
The effect of a quota changes $d\log y_{i^*}$ on output to a second order is

$$\Delta \log Y \approx \Pi_i d\log y_{i^*} + \frac{1}{2}\frac{d\Pi_i}{d\log y_{i^*}}(d\log y_{i^*})^2.$$

- Since first-order effect depends on profits, nonlinearities depend on change in profits.
- Does not depend on presence of other distortions, changes in other producers' profits.
- Can use input-output network, elasticities of substitution to compute $\frac{d\Pi_i}{d\log y_{i^*}}$.
  - Profits $=$ income of a "fixed factor," quota changes $=$ shocks to its productivity.
  - Apply existing results to calculate elasticity of profits to quota. (Baqaee and Farhi 2019).

Illustration: Nonlinearities in Horizontal Economy

- Response of profits to quota changes determines concavity / convexity of output:
  - Always concave around efficiency ($\Pi_i = 0$). May be convex away from frontier.



Special Case: Monopolist

- A tractable special case: Monopolist chooses output quota to maximize real profits.

Proposition (Nonlinearities with a Monopolist)

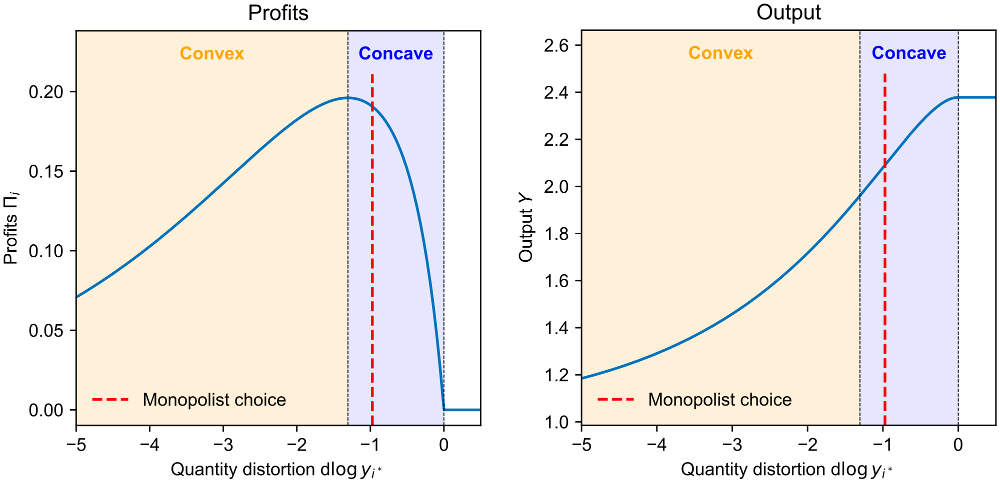
Suppose i produced by monopolist that chooses y_i to maximize real profits, taking all other quotas and production technologies as given. Then, effect of changes in the monopolist's quantity on output are

$$\Delta \log Y \approx \Pi_i d \log y_i - \frac{1}{2} \Pi_i^2 (d \log y_i)^2.$$

- Intuition: Monopolist's first-order condition determines response of profits to quantity.
- Profits sufficient to calculate how gains from increasing production peter out.

Illustration: Monopolist

- Monopolist always chooses quantity in concave region.



Nonlinearities: Multiple Quotas

- Method scales up to calculating nonlinear effects of multiple quota changes.

Proposition (Nonlinear Effects of Multiple Quotas)

The effect of a vector of quota changes $\mathbf{d}\log\mathbf{y}_*$ *on output to a second order is*

$$\Delta\log Y \approx \boldsymbol{\Pi}'\mathbf{d}\log\mathbf{y}_* + \frac{1}{2}(\mathbf{d}\log\mathbf{y}_*)'H(\mathbf{d}\log\mathbf{y}_*),$$

where H *is a symmetric matrix with* $H_{ij} = \partial\Pi_i/\partial\log y_{j^*}$.

- Quota demand system H summarizes responses of profits to all quotas.
- Since $\Pi_i = \partial\log Y/\partial\log y_{i^*}$, H is the Hessian of output w.r.t. quotas $\Rightarrow$ symmetric.

Illustrative Example: Multiple Quotas

- Consider horizontal economy. Response to changes in quotas on firms 1 and 2:

$$\Delta \log Y \approx \underbrace{\Pi_1 d\log y_{1^*} + \Pi_2 d\log y_{2^*}}_{\text{First order}}$$
$$+ \underbrace{(1/2)\left(H_{11}\left(d\log y_{1^*}\right)^2 + H_{22}\left(d\log y_{2^*}\right)^2 + 2H_{12}\left(d\log y_{1^*}\right)\left(d\log y_{2^*}\right)\right)}_{\text{Second order}}.$$

- H_{12} determines whether relaxing y_{1^*} amplifies/reduces gains from relaxing y_{2^*}.

- In horizontal economy, $H_{12} > 0$ if

$$\theta < 1 - \frac{(\lambda_1 - \Pi_1)(\lambda_2 - \Pi_2)}{(1-\lambda_1-\lambda_2)\Pi_1\Pi_2}.$$

Necessary condition is that $\theta < 1$, i.e., outputs of firms 1 and 2 are complements.

A Second Special Case: Distance to the Efficient Frontier

Proposition (Distance to the Frontier)

Suppose we remove a subset of quotas $\mathscr{I}^*$. Let $\mathbf{y}^{\mathbf{eff}}$ be the output quantities that result if all quotas in $\mathscr{I}^*$ are removed, and define $\mathbf{d}\log\mathbf{y}_* \equiv \log\mathbf{y}_* - \log\mathbf{y}^{\mathbf{eff}}$.

The gains from removing all quotas in $\mathscr{I}^*$ are

$$\Delta\log Y \approx -\frac{1}{2}\left(\mathbf{d}\log\mathbf{y}_*\right)' H\left(\mathbf{d}\log\mathbf{y}_*\right), \tag{Option 1}$$

Alternatively,

$$\Delta\log Y \approx -\frac{1}{2}\mathbf{\Pi}'\mathbf{d}\log\mathbf{y}_*. \tag{Option 2}$$

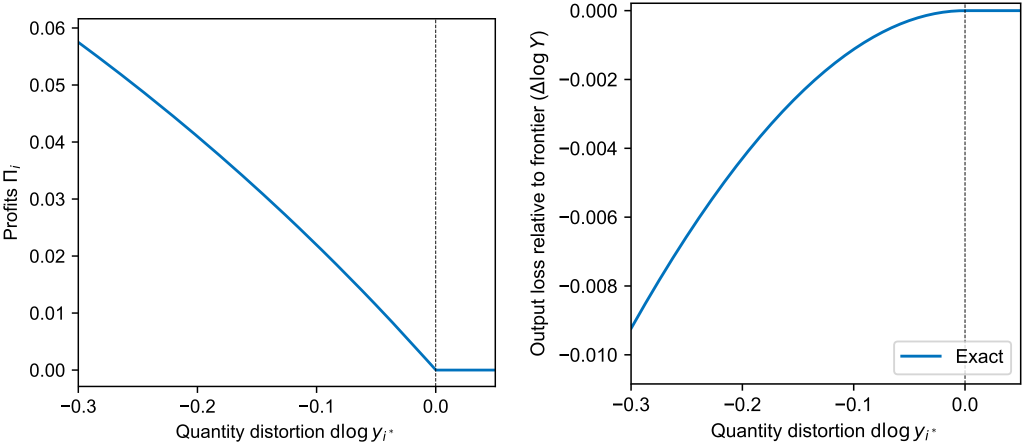
Equivalently,

$$\Delta\log Y \approx -\frac{1}{2}\mathbf{\Pi}' H^{-1}\mathbf{\Pi}. \tag{Option 3}$$

The approximation error for each expression is order $\left(d\log y_*\right)^3$.

Distance to Frontier Example: Round-about Economy

- Round-about economy: Firm uses labor + its own output as intermediate input.
- Quota on use of intermediate input.



Distance to the Frontier: Option 1

Corollary (Option 1)

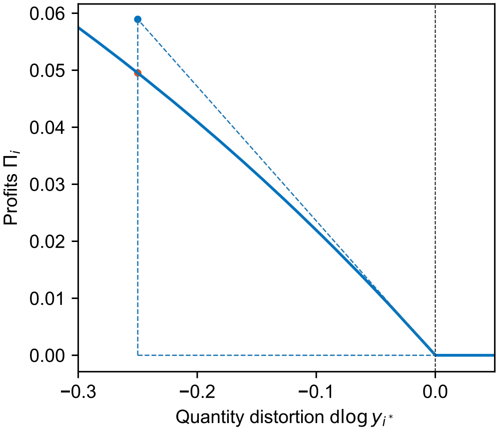
The distance to the frontier to a second order is

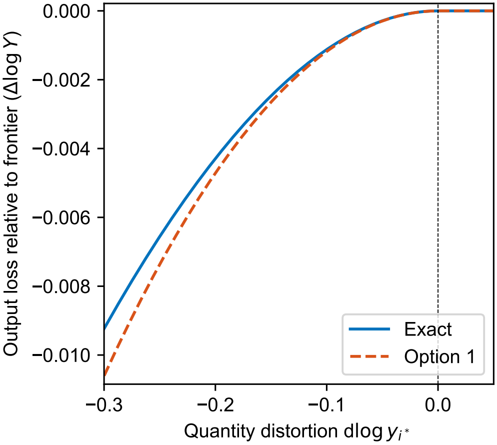
$$\Delta \log Y \approx -\frac{1}{2}\frac{\partial \Pi_i}{\partial \log y_{i^*}}(d \log y_{i^*})^2.$$

- Special case of nonlinearities: starting at efficiency, profits $\Pi_i = 0$.
- Can compute semi-elasticity $\partial \Pi_i / \partial \log y_{i^*}$ in terms of input-output matrix, elasticities of substitution.

Distance to Frontier Example: Option 1

- Option 1: $\Delta \log Y \approx 1/2 \, \frac{\partial \Pi_i}{\partial \log y_{i^*}} (d \log y_{i^*})^2 = -\frac{1}{2\theta_1} \frac{\lambda_1 - 1}{\lambda_1} (d \log y_{i^*})^2.$





Distance to the Frontier: Option 2

- Option 1 extrapolates profits from the efficient point using H.
- Option 2: Use profits at distorted allocation: $\frac{\partial \Pi_i}{d \log y_{i^*}} d \log y_{i^*} \approx \Pi_i$.

Corollary (Option 2)

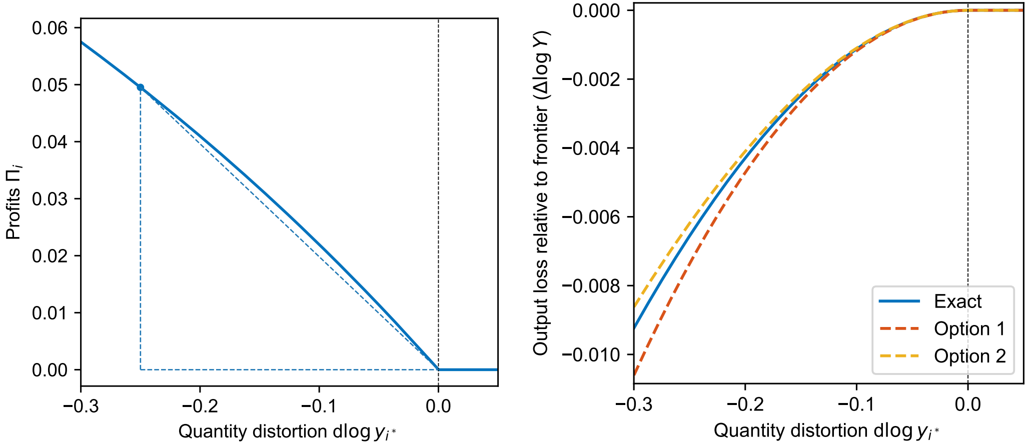
The distance to the frontier to a second order is

$$\Delta \log Y \approx -\frac{1}{2} \Pi_i d \log y_{i^*}.$$

- I.e., estimate distance to frontier using $-1/2 \times$ **profits** $\times$ **size of distortion**.
- Intuition: Average of first-order at inefficient point ($\Pi_i \, d \log y_{i^*}$) and at efficient point (0).

Distance to Frontier Example: Option 2

- Option 2: $\Delta \log Y \approx -1/2 \times$ **profits** $\times$ **size of distortion** $= -1/2\, \Pi_i d \log y_{i^*}$.



Distance to the Frontier

- Both formulas require knowing the efficient level of output $d \log y_{i^*} = \log y_{i^*} - \log y_i^{\text{eff}}$.
- Option 3: Estimate distance using **response of profits to quota changes**.

Corollary (Option 3)

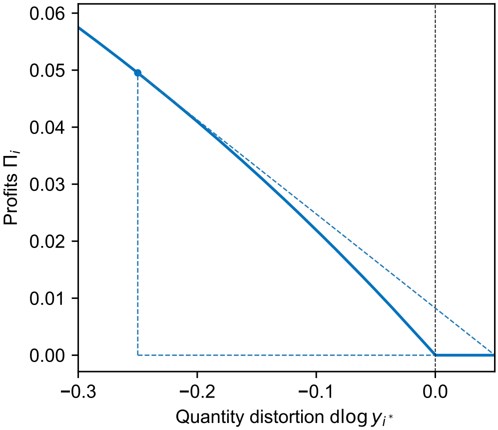
The distance to the frontier to a second order is

$$\Delta \log Y \approx -\frac{1}{2}\mathbf{\Pi}' H^{-1} \mathbf{\Pi} = \frac{1}{2}\Pi_i \left[-\frac{d \log \Pi_i}{d \log y_{i^*}} \right]^{-1}.$$

- Intuition: If profits fall quickly with output, close to efficient level $\Rightarrow$ smaller gains.

Distance to Frontier Example: Option 3

- Option 3: Use elasticity of profits to quota: $\Delta \log Y \approx 1/2\, \Pi_i [-\frac{d \log \Pi_i}{d \log y_{i^*}}]^{-1}$.



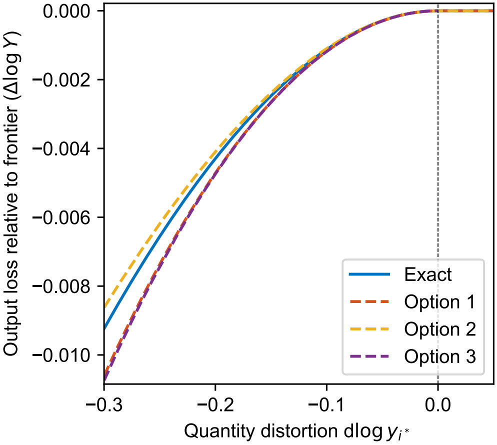


Table of Contents

Empirical Examples

1. Distance to Frontier: Argentina Capital Controls.
2. Quota Demand System: Quotas on Chinese Textile & Clothing Exports.
3. Nonlinearities Far from the Frontier: NYC Taxicab Medallions.

Distance to Frontier: Argentina's Capital Controls

- In September 2019, Argentina reimposed capital controls, restricting amount that households and firms could access foreign exchange markets.
- Restrictions led to decline in capital outflows from \$7.2B to $<$\$1.5B per month.
- What are the efficiency losses from these restrictions?
  - Permitted transactions take place at official exchange rate $\bar{e}$.
  - Unconstrained transactions take place at black-market exchange rate e.
  - Gap between $\bar{e}$ and e are profits earned by permit to exchange under controls.

Distance to Frontier: Argentina's Capital Controls

- Applying "Option 2" for distance to the frontier:

$$\Delta \log Y \approx -\frac{1}{2}\Pi_i d\log y_{i^*} \approx -\frac{1}{2}\left(\log e/\bar{e}\right) dy_{i^*}.$$

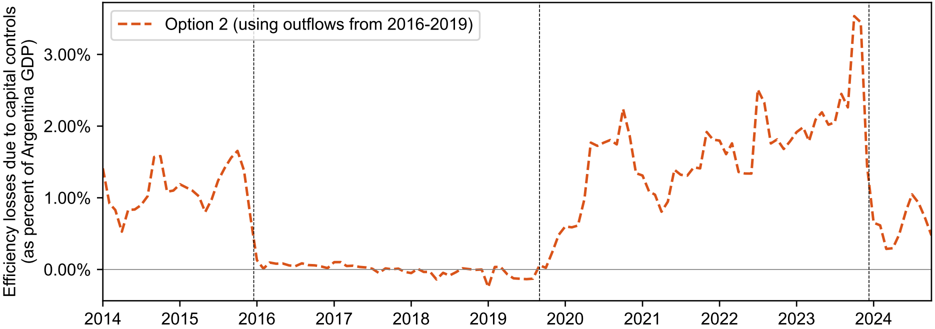
- Measure distortion dy_{i^*} as gap relative to 2016–2019 outflows.

Distance to Frontier: Argentina's Capital Controls

- Applying "Option 2" for distance to the frontier:

$$\Delta \log Y \approx -\frac{1}{2}\Pi_i d \log y_{i^*} \approx -\frac{1}{2}\left(\log e/\bar{e}\right) dy_{i^*}.$$

- Measure distortion dy_{i^*} as gap relative to 2016–2019 outflows.



Vertical lines indicate end of capital controls (Dec 17 2015), reinstatement of capital controls (Sep 1 2019), and devaluation of peso (Dec 10 2023).

Distance to Frontier: Argentina's Capital Controls

- *Currency elasticity* θ: Depreciation in nominal exchange rates caused by purchases of foreign currency equal to one percent of GDP. (Blanchard et al. 2015; Adler et al. 2019).

- "Option 3" uses θ to estimate size of distortion dy_{i^*}:

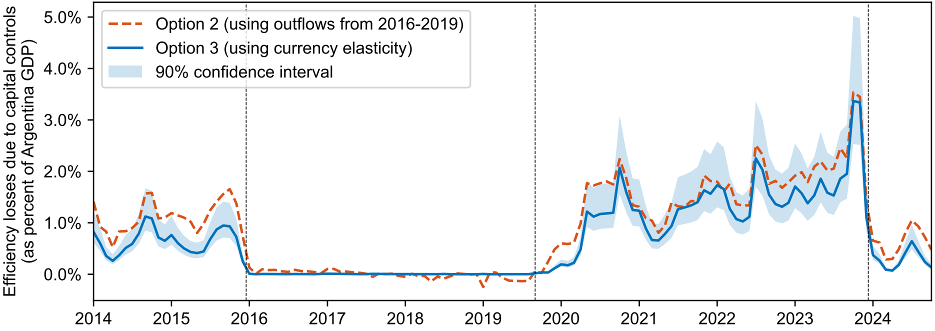
$$\theta = \frac{\log e/\bar{e}}{dy_{i^*}/\text{GDP}} \qquad \Rightarrow \qquad \frac{\Delta Y}{\text{GDP}} \approx -\frac{1}{2}\frac{1}{\theta}\left(\log e/\bar{e}\right)^2 .$$

Distance to Frontier: Argentina's Capital Controls

- *Currency elasticity* θ: Depreciation in nominal exchange rates caused by purchases of foreign currency equal to one percent of GDP. (Blanchard et al. 2015; Adler et al. 2019).

- "Option 3" uses θ to estimate size of distortion dy_{i^*}:

$$\theta = \frac{\log e/\bar{e}}{dy_{i^*}/\text{GDP}} \qquad \Rightarrow \qquad \frac{\Delta Y}{\text{GDP}} \approx -\frac{1}{2}\frac{1}{\theta}\left(\log e/\bar{e}\right)^2 .$$



Vertical lines indicate end of capital controls (Dec 17 2015), reinstatement of capital controls (Sep 1 2019), and devaluation of peso (Dec 10 2023).

Quota Demand System: China's Textile and Clothing Exports

- 1975–1994 Multi-Fiber Agreement capped China's textile & clothing exports to US, EU.
- Agreement to phase-out quotas from 1995–2005 as part of WTO Uruguay Round.
- Staged phase-out allows us to study interaction between quotas.
  - Most quotas removed in Jan 2002 (Phase III) and Jan 2005 (Phase IV).
  - Phase III: Knit fabrics, gloves, dressing gowns, brassieres, and textile luggage products.
  - Phase IV: Silk, wool, and cotton textiles, carpets, most apparel categories.

Quota Demand System: China's Textile and Clothing Exports

1. Use quota auction prices to estimate initial quota profits.
2. Use response of export quantities to recover quota demand system, H.
3. Use quota demand system to estimate gains from relaxing any subset of quotas.

Quota Demand System: China's Textile and Clothing Exports

1. Use quota auction prices to estimate initial quota profits.
   - Market prices for quotas to export in each product category in 2001.
   - Quota profits for Phase III and Phase IV goods, $\Pi_{\text{Phase III}} = \$38\text{B}$, $\Pi_{\text{Phase IV}} = \$394\text{B}$.
2. Use response of export quantities to recover quota demand system, H.
3. Use quota demand system to estimate gains from relaxing any subset of quotas.

Quota Demand System: China's Textile and Clothing Exports

1. Use quota auction prices to estimate initial quota profits.

2. Use response of export quantities to recover quota demand system, H.

   - Because profits go to zero when quotas are removed, H solves:

$$
\begin{aligned}
\Pi_{\text{Phase III}} &= \left(d\log y^{\text{III}}_{\text{Phase III}}\right) H_{11}\,, \\
\Pi_{\text{Phase III}} &= \left(d\log y^{\text{IV}}_{\text{Phase III}}\right) H_{11} + \left(d\log y^{\text{IV}}_{\text{Phase IV}}\right) H_{12}\,, \\
\Pi_{\text{Phase IV}} &= \left(d\log y^{\text{III}}_{\text{Phase IV}}\right) H_{21} + \left(d\log y^{\text{IV}}_{\text{Phase IV}}\right) H_{22}\,,
\end{aligned}
$$

   where $d\log y^{\text{III}}_{x}$ ($d\log y^{\text{IV}}_{x}$) is the change in exports for x goods following Phase III (IV).

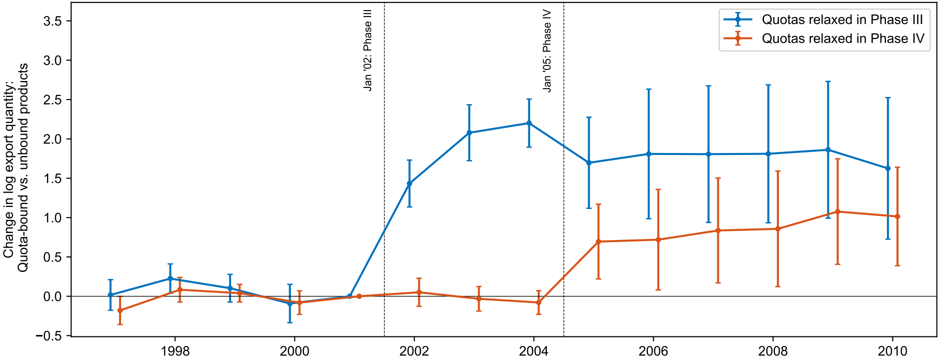
   - Symmetry implies $H_{12} = H_{21}$.

3. Use quota demand system to estimate gains from relaxing any subset of quotas.

Quota Demand System: China's Textile and Clothing Exports

$$\log y_{ict} = \beta_t^{\text{Phase III}} (\text{Binding}_c \times 1\{c \text{ quota relaxed in Phase III}\} \times 1\{\text{year} = t\})$$
$$+ \beta_t^{\text{Phase IV}} (\text{Binding}_c \times 1\{c \text{ quota relaxed in Phase IV}\} \times 1\{\text{year} = t\}) + \alpha_t + \delta_i + \varepsilon_{ict},$$

- E.g., $\beta_t^{\text{Phase III}}$ is change in Phase III good exports in t relative to unconstrained goods.



Quota Demand System: China's Textile and Clothing Exports

1. Use quota auction prices to estimate initial quota profits.

2. Use response of export quantities to recover quota demand system, H.

   - Because profits go to zero when quotas are removed, H solves:

$$\begin{aligned}
\Pi_{\text{Phase III}} &= (d\log y^{\text{III}}_{\text{Phase III}})\, H_{11}\,, \\
\Pi_{\text{Phase III}} &= (d\log y^{\text{IV}}_{\text{Phase III}})\, H_{11} + (d\log y^{\text{IV}}_{\text{Phase IV}})\, H_{12}\,, \\
\Pi_{\text{Phase IV}} &= (d\log y^{\text{III}}_{\text{Phase IV}})\, H_{21} + (d\log y^{\text{IV}}_{\text{Phase IV}})\, H_{22}\,,
\end{aligned}$$

   where $d\log y^{\text{III}}_{x}$ ($d\log y^{\text{IV}}_{x}$) is the change in exports for x goods following Phase III (IV).

   - Symmetry implies $H_{12} = H_{21}$.

3. Use quota demand system to estimate gains from relaxing any subset of quotas.

Quota Demand System: China's Textile and Clothing Exports

1. Use quota auction prices to estimate initial quota profits.
2. Use response of export quantities to recover quota demand system, H.
3. Use quota demand system to estimate gains from relaxing any subset of quotas.
   - Estimated inverse quota demand system:

$$\Pi = \begin{bmatrix} \Pi_{\text{Phase III}} \\ \Pi_{\text{Phase IV}} \end{bmatrix} = \begin{bmatrix} \$38\text{B} \\ \$394\text{B} \end{bmatrix}, \qquad H = \begin{bmatrix} -17.9 & -7.6 \\ -7.6 & -495.6 \end{bmatrix}.$$

Quota Demand System: China's Textile and Clothing Exports

1. Use quota auction prices to estimate initial quota profits.

2. Use response of export quantities to recover quota demand system, H.

3. Use quota demand system to estimate gains from relaxing any subset of quotas.
   - Estimated inverse quota demand system:

$$\Pi = \begin{bmatrix} \Pi_{\text{Phase III}} \\ \Pi_{\text{Phase IV}} \end{bmatrix} = \begin{bmatrix} \$38\text{B} \\ \$394\text{B} \end{bmatrix}, \qquad H = \begin{bmatrix} -17.9 & -7.6 \\ -7.6 & -495.6 \end{bmatrix}.$$

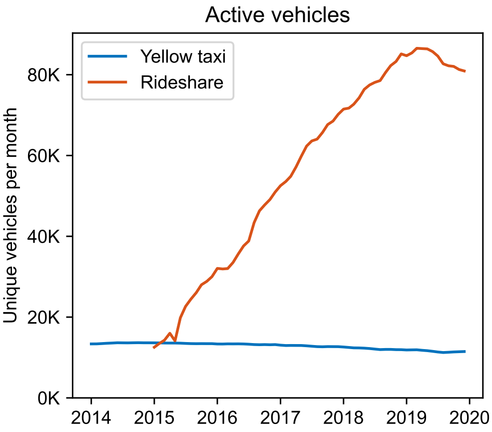
Intervention	Efficiency gains (2001 USD $B)
(A) Relaxing Phase III quotas only	$40
(B) Relaxing Phase IV quotas only	$158
(C) Relaxing both Phase III and IV quotas	$185
Difference: $\mathrm{C} - (\mathrm{A} + \mathrm{B})$	$13

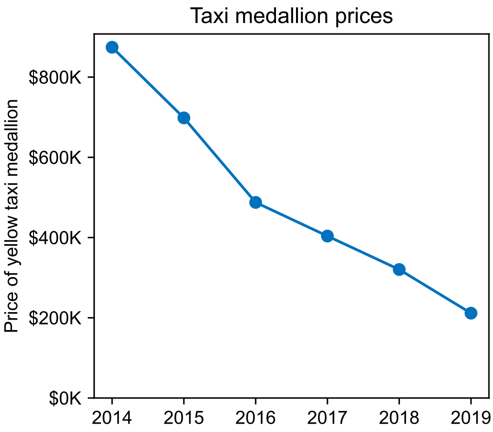
Nonlinearities: Taxicab Medallions

- Since 1937, quota on NYC taxicab medallions restricting total supply to $\approx$ 14k.

Nonlinearities: Taxicab Medallions

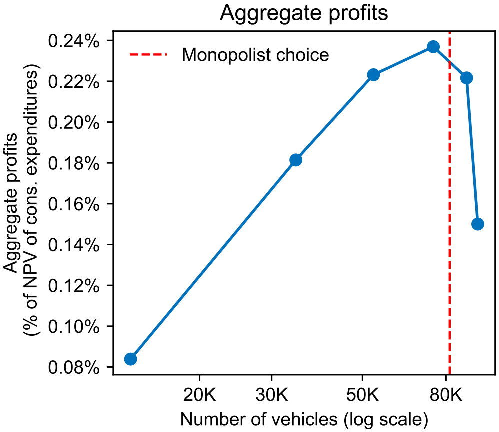
- Since 1937, quota on NYC taxicab medallions restricting total supply to $\approx$ 14k.
- Use arrival of rideshare apps in NYC to quantify gains from relaxing quota on cabs.

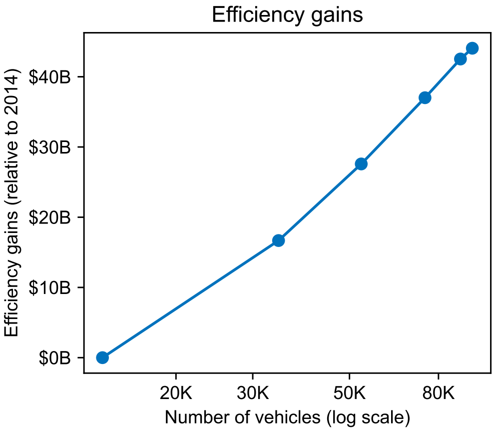




Nonlinearities: Taxicab Medallions

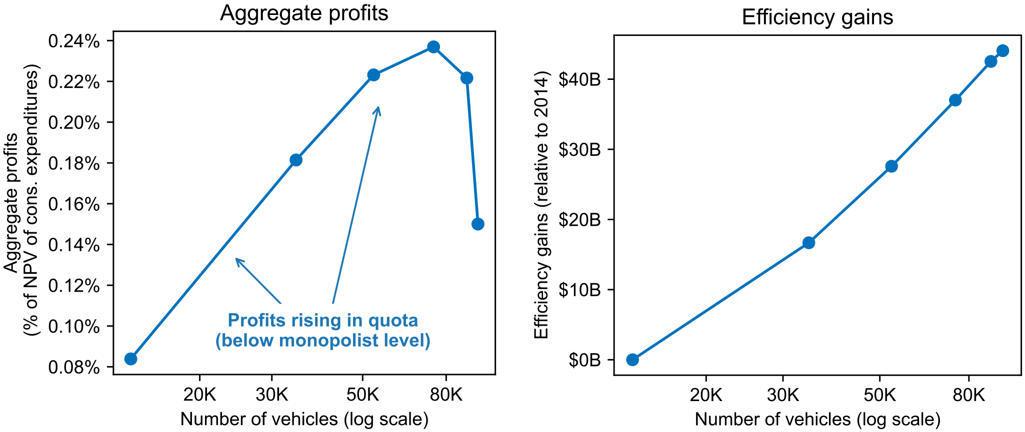
- Assume that medallion transaction prices reflect rents accruing to owners.
- Gains from relaxing taxicab quota are $\Delta \log Y_t \approx \left(\Pi_{it} + \frac{1}{2} d\Pi_{it}\right) d \log y_{i^* t}$.





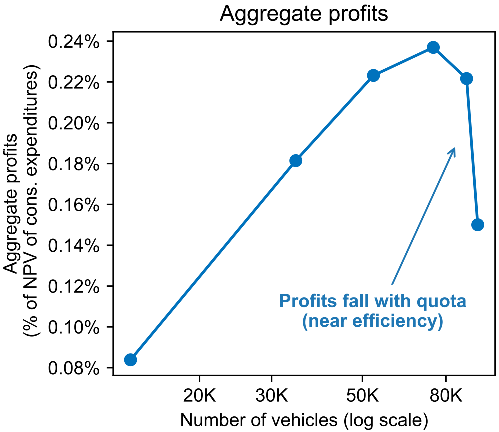
Nonlinearities: Taxicab Medallions

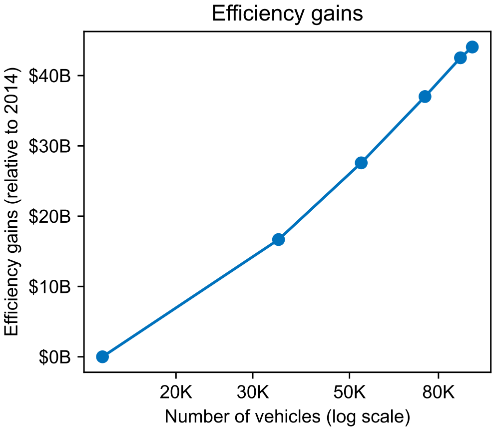
- Assume that medallion transaction prices reflect rents accruing to owners.
- Gains from relaxing taxicab quota are $\Delta \log Y_t \approx \left(\Pi_{it} + \frac{1}{2} d\Pi_{it}\right) d \log y_{i^* t}$.



Nonlinearities: Taxicab Medallions

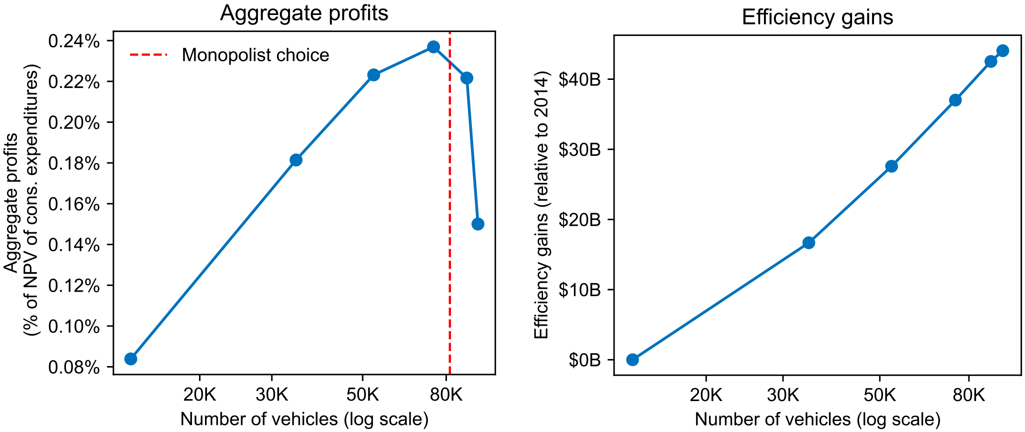
- Assume that medallion transaction prices reflect rents accruing to owners.
- Gains from relaxing taxicab quota are $\Delta \log Y_t \approx \left(\Pi_{it} + \frac{1}{2} d\Pi_{it}\right) d \log y_{i^* t}$.





Nonlinearities: Taxicab Medallions

- Assume that medallion transaction prices reflect rents accruing to owners.
- Gains from relaxing taxicab quota are $\Delta \log Y_t \approx \left(\Pi_{it} + \frac{1}{2} d\Pi_{it}\right) d \log y_{i^* t}$.



Nonlinearities: Taxicab Medallions

- Gains from relaxing quota over 2014–2019.
  - Cumulating gains over each year: $\Delta \log Y \approx \sum_t \left(\Pi_{it} + \frac{1}{2} d\Pi_{it}\right) d \log y_{i^* t}$.
- Not efficient at the end. What is the remaining distance to frontier?
  - Use elasticity of profits to quantity in final year: $\Delta \log Y \approx \frac{1}{2}\Pi_i \left[-\frac{d \log \Pi_i}{d \log y_{i^*}}\right]^{-1}$.

	Change from 2014–2019	Distance to frontier
Output gains	\$44.1B	\$1.8B
Gains per New York MSA household	\$6,029	\$246
% of NPV of transportation expenditures (incl. vehicles/gas)	2.61%	0.11%

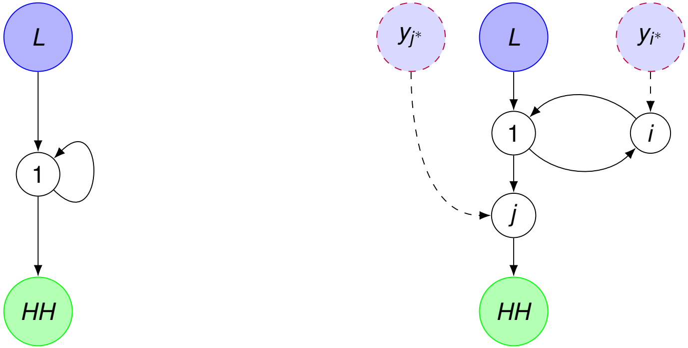
Conclusion

- General framework for analyzing economies with quota distortions.
- Two lessons:
  - 1. Any distorted allocation can be recast as equilibrium of an economy with quotas.
  - 2. Economies with quotas are constrained efficient, and thus highly tractable.
- Effects of quota changes, productivity shocks disciplined by a few sufficient stats.
- Nonlinearities, distance to efficient frontier using inverse quota demand system.
- Examples of how to apply results (H-1B caps, zoning laws, capital controls, export quotas, taxicabs).

Table of Contents

Implementing an Allocation Using Quotas: Example

- Round-about economy. Feasible allocations: $\{(y_1, c_1, x_{11}) \mid c_1 + x_{11} \leq y_1 = F_1(L, x_{11})\}$.



(a) Original.

(b) With quotas.

- In fact, more general than wedges: can implement allocation when L, x_{11} are perfect substitutes / complements.

Matching observables in wedge and quota economies

- A given allocation can typically be implemented by many different sets of wedges.
- $\Rightarrow$ Same allocation implemented with wedges vs. quotas can differ in prices/profits.

Proposition (Matching observables in wedge and quota economies)

Consider an economy with quotas in which all producer prices p_i and factor wages w_f are strictly positive. Consider a second economy in which the same allocation of resources is implemented with wedges, τ.

If $\tau_i \geq 1$ for all i and, for each good or factor i, either the good is directly consumed by the household $c_i > 0$ or there exists some producer j such that $\partial F_j / \partial x_{ji} > 0$ and $\tau_j = 1$, then prices, sales, and profits are identical across the two economies.

- Intuition: Profits measured relative to one unconstrained user of each resource.

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica 80*(5), 1977–2016.

Adler, G., N. Lisack, and R. C. Mano (2019). Unveiling the effects of foreign exchange intervention: A panel approach. *Emerging Markets Review* (100620).

Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal: Macroeconomics 9*(4), 254–280.

Baqaee, D. R. (2018). Cascading failures in production networks. *Econometrica 86*(5), 1819–1838.

Baqaee, D. R. and E. Farhi (2019). The macroeconomic impact of microeconomic shocks: Beyond Hulten's theorem. *Econometrica 87*(4), 1155–1203.

Baqaee, D. R. and E. Farhi (2020). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics 135*(1), 105–163.

Basu, S. and J. G. Fernald (2002). Aggregate productivity and aggregate technology. *European Economic Review 46*(6), 963–991.

Bigio, S. and J. La'O (2020). Distortions in production networks. *The Quarterly Journal of Economics 135*(4), 2187–2253.

Blanchard, O., G. Adler, and I. de Carvalho Filho (2015). Can foreign exchange intervention stem exchange rate pressures from global capital flow shocks. IMF Working Paper.

Brambilla, I., A. K. Khandelwal, and P. K. Schott (2010). *China's Growing Role in World Trade*, Chapter China's Experience under the Multi-Fiber Arrangement (MFA) and the Agreement on Textiles and Clothing (ATC), pp. 345–387. University of Chicago Press.

Buera, F. J. and N. Trachter (2024). Sectoral development multipliers. Technical Report 32230, National Bureau of Economic Research.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2007). Business cycle accounting. *Econometrica 75*(3), 781–836.

Clemens, M. A. (2013). Why do programmers earn more in Houston than Hyderabad? Evidence from randomized processing of US visas. *American Economic Review 103*(3), 198–202.

Dasgupta, P. and J. E. Stiglitz (1977). Tariffs vs. quotas as revenue raising devices under uncertainty. *American Economic Review 67*(5), 975–981.

Domar, E. D. (1961). On the measurement of technological change. *The Economic Journal 71*(284), 709–729.

Edmond, C., V. Midrigan, and D. Y. Xu (2023). How costly are markups? *Journal of Political Economy 131*(7), 1619–1675.

Feenstra, R. C. (1988). Quality change under trade restraints in Japanese autos. *The Quarterly Journal of Economics 103*(1), 131–146.

Feenstra, R. C. (1992). How costly is protectionism? *Journal of Economic Perspectives 6*(3), 159–178.

Foerster, A. T., P.-D. Sarte, and M. W. Watson (2011). Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy 119*(1), 1–38.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*(3), 733–772.

Glaeser, E. and J. Gyourko (2018). The economic implications of housing supply. *Journal of Economic Perspectives 32*(1), 3–30.

Grassi, B. (2017). IO in I-O: Size, industrial organization, and the input-output network make a firm structurally important. Working paper.

Gyourko, J. and J. Krimmel (2021). The impact of local residential land use restrictions on land values across and within single family housing markets. *Journal of Urban Economics 126*, 103374.

Harberger, A. C. (1954). Monopoly and resource allocation. *American Economic Review 44*(2), 77–87.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Hsieh, C.-T. and E. Moretti (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics 11*(2), 1–39.

Hulten, C. R. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies 45*(3), 511–518.

Khandelwal, A. K., P. K. Schott, and S.-J. Wei (2013). Trade liberalization and embedded institutional reform: Evidence from Chinese exporters. *American Economic Review 103*(6), 2169–2195.

Liu, E. (2019). Industrial policies in production networks. *The Quarterly Journal of Economics 134*(4), 1883–1948.

Petrin, A. and J. Levinsohn (2012). Measuring aggregate productivity growth using plant-level data. *The RAND Journal of Economics 43*(4), 705–725.

Reischer, M. (2019). Finance-thy-neighbor: Trade credit origins of aggregate fluctuations. Working paper.

Restuccia, D. and R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics 11*(4), 707–720.

Weitzman, M. L. (1974). Prices vs. quantities. *The Review of Economic Studies 41*(4), 477–491.