Everything All the Time?
Entry and Exit in U.S. Import Varieties

Roc Armenter (FRB Philadelphia)
Miklós Koren (CEU and CEPR)

St.Louis Fed/Wash U. - September 2013

Introduction

Introduction

- How much have newly-traded goods contributed to overall trade growth?

Introduction

- How much have newly-traded goods contributed to overall trade growth?
- Why the extensive margin warrants our attention:
  - Trade benefits depend on it: consumer surplus from new varieties, pro-competitive effects, and aggregate TFP gains.
  - The extensive margin may have its own idiosyncratic response to price or economic fluctuations.
  - Implications for the evaluation and design of trade policies.

Introduction

- How much have newly-traded goods contributed to overall trade growth?
- Why the extensive margin warrants our attention:
  - Trade benefits depend on it: consumer surplus from new varieties, pro-competitive effects, and aggregate TFP gains.
  - The extensive margin may have its own idiosyncratic response to price or economic fluctuations.
  - Implications for the evaluation and design of trade policies.
- We propose a new theory of the extensive margin.

Our model

- We introduce a novel demand system based on a standard random-utility, discrete choice model of product differentiation.

Our model

- We introduce a novel demand system based on a standard random-utility, discrete choice model of product differentiation.
- Whereas traditional demand systems predict market shares,

Our model

- We introduce a novel demand system based on a standard random-utility, discrete choice model of product differentiation.
- Whereas traditional demand systems predict market shares,
- ...our model yields instead the probability that a purchase of a given good is supplied by any given country.

Our model

- We introduce a novel demand system based on a standard random-utility, discrete choice model of product differentiation.
- Whereas traditional demand systems predict market shares,
- ...our model yields instead the probability that a purchase of a given good is supplied by any given country.
- Crucially, there are only a finite number of independent purchase decisions per period.

Our model

- We introduce a novel demand system based on a standard random-utility, discrete choice model of product differentiation.
- Whereas traditional demand systems predict market shares,
- ...our model yields instead the probability that a purchase of a given good is supplied by any given country.
- Crucially, there are only a finite number of independent purchase decisions per period.
- The demand of any variety of any good is a random variable.

The extensive margin

- There is a positive probability that a commodity is not traded during a period.
  - The exact probability will be determined by the price vector and structural parameters.

The extensive margin

- There is a positive probability that a commodity is not traded during a period.
  - The exact probability will be determined by the price vector and structural parameters.
- There is a rich set of predictions regarding the extensive and intensive margin both in the cross-section and across time.

The extensive margin

- There is a positive probability that a commodity is not traded during a period.
  - The exact probability will be determined by the price vector and structural parameters.
- There is a rich set of predictions regarding the extensive and intensive margin both in the cross-section and across time.
- The model naturally reconciles two key observations:
  1. Cross-section: There are many missing trade flows,
  2. Across time: There is a lot of churning of goods.

  through the law of rare events.

Going quantitative

- The distribution of random-utility terms is given by a Gumbel distribution which is key to keep the model tractable despite the underlying probabilistic structure.
- As the number of purchases tends to infinite, our model collapses to the standard CES demand system.
  - As first pointed out by Anderson et al. (1987).
- We calibrate the model to U.S. imports data 1990-2001 at HS10 product - country level.
- The quantitative performance of the model is very good.

Brief literature review

- Armenter and Koren (2010) establish trade data are typically sparse, i.e., too few observations given the number of categories.
  - Propose an atheoretical benchmark to account for sparsity, and find that several stylized facts are matched pretty well.
  - It is hard to identify the relevant theory of the extensive margin.
- Eaton, Kortum, and Sotelo (2012) consider a finite number of firms in a standard model.
  - Supply and demand granularity can be complementary.
  - We have independent realizations per period: dynamic facts.
- Some other work generating sparse trade data:
  - Kropf and Saure (2011) and Hornok and Koren (2012) assume fixed costs per shipment.
  - Bekes and Murakozy (2012) tackle temporary trade at the firm level.

Outline

1. Model,
2. Data and calibration,
3. Cross-section results,
4. Dynamic results,
5. The extensive margin and NAFTA,
6. Re-evaluating welfare gains from new varieties,
7. Conclusions.

The Model

Commodity space and purchase decisions

- There are J countries each supplying a differentiated variety of each product (or good) $g \in G$.

Commodity space and purchase decisions

- There are J countries each supplying a differentiated variety of each product (or good) $g \in G$.
- For each period $t = 1, 2, \ldots$ there is a finite number of purchase decisions (or transactions) for each good $g \in G$,

$$n_t^g = \lceil \alpha^g n_t \rceil \in \mathbb{N}$$

  where n_t is overall trade intensity and $\alpha^g > 0, \sum_G \alpha^g = 1$.
- Each purchase is endowed with a budget $y_t = Y_t / n_t$.
- Aggregate income Y_t and purchase decisions n_t grow at constant rates γ_y and γ_n.

Discrete choice, random utility model

- Each purchase decision is an independent, discrete choice between each of the country varieties $j = 1, 2, \ldots, J$.

Discrete choice, random utility model

- Each purchase decision is an independent, discrete choice between each of the country varieties $j = 1, 2, \ldots, J$.
- It is a *discrete* choice because each purchase decision must be satisfied by a single country—albeit in the quantity of choice.

Discrete choice, random utility model

- Each purchase decision is an independent, discrete choice between each of the country varieties $j = 1, 2, \ldots, J$.
- It is a *discrete* choice because each purchase decision must be satisfied by a single country—albeit in the quantity of choice.
- It is an *independent* choice each purchase for good g has its own idiosyncratic type $\theta \in \Re_+^J$, drawn from distribution F^g independently of other purchases in any good or period.

Discrete choice, random utility model

- Preferences over varieties are given by

$$u\left(x, c; \theta\right) = \left(\sum_{J} \theta_j x_j\right)^{\eta} c^{1-\eta},$$

  - x_j is the individual's purchase of each variety,
  - c is a non-traded numeraire good.

Discrete choice, random utility model

- Preferences over varieties are given by

$$u(x, c; \theta) = \left(\sum_J \theta_j x_j \right)^\eta c^{1-\eta},$$

  - x_j is the individual's purchase of each variety,
  - c is a non-traded numeraire good.
- Indirect utility function,

$$v_{jt}^g = \max_{x_j, c} \{ \eta \ln x_j + (1-\eta) \ln c : p_{jt}^g x_j + c \leq y_t \}.$$

- The numeraire good and elasticity η are relevant only for welfare analysis.

Discrete choice, random utility model

- The optimal choice by a consumer of type θ solves

$$V_t^g\left(p_t^g;\theta\right) = \max\{v_t(p_{jt}^g) + \ln\theta_j : j \in \{1, 2, \ldots, J\}\}.$$

- Prices plus preferences determine the variety of choice.
- Purchase-level demand of variety j:

$$x_{jt}^g = \frac{\eta y_t}{p_{jt}^g}$$

Random utility terms

- Let $\ln \theta_j$ be i.i.d. according to a Gumbel (or type I extreme value) distribution,

$$F^g(m) = \exp\left(-\exp\left(-\frac{m}{\mu^g}\right)\right).$$

- Parameter $\mu^g > 0$ governs the dispersion of θ and the sensitivity to prices.

Random utility terms

- Let $\ln \theta_j$ be i.i.d. according to a Gumbel (or type I extreme value) distribution,

$$F^g(m) = \exp\left(-\exp\left(-\frac{m}{\mu^g}\right)\right).$$

- Parameter $\mu^g > 0$ governs the dispersion of θ and the sensitivity to prices.
- Independence, independence, independence
  - Across varieties, goods, purchases, and time.

Choice probability

- The Gumbel distribution keeps the demand system tractable due to its properties regarding the maximum order statistic.
- Given $p_t^g = \left\{ p_{jt}^g : j = 1, 2, \ldots, J \right\}$, the probability that country i supplies any given purchase decision for good g is

$$s_{it}^g = \frac{\left(p_{it}^g\right)^{-1/\mu^g}}{\sum_J \left(p_{jt}^g\right)^{-1/\mu^g}}.$$

Retrieving the CES demand as $n_t = \infty$

- Let the number of purchases per period be $n_t = \infty$.
- By the law of large numbers, s_{jt}^{g} converges to the fraction of purchase decisions provided by country j, the market share.

Retrieving the CES demand as $n_t = \infty$

- Let the number of purchases per period be $n_t = \infty$.
- By the law of large numbers, s^g_{jt} converges to the fraction of purchase decisions provided by country j, the market share.
- Letting $\rho^g = \frac{1}{1+\mu^g}$,

$$X^g_{jt} = \left(\frac{p^g_{jt}}{P^g_t}\right)^{-\frac{1}{1-\rho^g}} \frac{Y^g_t}{P^g_t}$$

where

$$P^g_t = \left(\sum_J \left(p^g_{jt}\right)^{-\frac{\rho^g}{1-\rho^g}}\right)^{-\frac{1-\rho^g}{\rho^g}}.$$

- This result is due to Anderson, Palma, and Thisse (EL, 1987).

Everything all the time as $n_t = \infty$

- If a variety is available, it is traded $X_{jt}^{g} > 0$ at all times.

Everything all the time as $n_t = \infty$

- If a variety is available, it is traded $X^g_{jt} > 0$ at all times.
- Trade is akin to oil flowing through a pipeline:
  - Demand scales proportionally with the period length.
  - We should observe the same set of traded goods for any time sub-interval.

Everything all the time as $n_t = \infty$

- If a variety is available, it is traded $X^g_{jt} > 0$ at all times.
- Trade is akin to oil flowing through a pipeline:
  - Demand scales proportionally with the period length.
  - We should observe the same set of traded goods for any time sub-interval.
- This property is the key identification step for Feenstra (1994) and Broda and Weinstein (2006).
- It also features in countless calibration schemes.

CES connnections

- The elasticity of substitution relates to the dispersion parameter as
$$\sigma^g = \frac{1 + \mu^g}{\mu^g}.$$
- High elasticity when there is little dispersion in θ: price is all that matters, few idiosyncratic tastes.

CES connnections

- The elasticity of substitution relates to the dispersion parameter as
$$\sigma^g = \frac{1+\mu^g}{\mu^g}.$$
- High elasticity when there is little dispersion in θ: price is all that matters, few idiosyncratic tastes.
- "Love of variety:" More goods mean more chances to get a high θ_j — maximum V_t^g is expected to grow with the number of "draws."
  - High dispersion makes varieties more valuable.

Back to our model: finite n_t

- Let z_{jt}^g be the number of purchases of variety j in good g.
- The vector $z_t^g = \{z_{jt}^g : j = 1, \ldots, J\}$ is distributed according to a Multinomial

$$\Pr(z) = \frac{n_t^g!}{z_1! z_2! \ldots z_J!} \left(s_{1t}^g\right)^{z_1} \left(s_{2t}^g\right)^{z_2} \ldots \left(s_{Jt}^g\right)^{z_J}.$$

Back to our model: finite n_t

- Let z^g_{jt} be the number of purchases of variety j in good g.
- The vector $z^g_t = \{z^g_{jt} : j = 1, \ldots, J\}$ is distributed according to a Multinomial

$$\Pr(z) = \frac{n^g_t!}{z_1! z_2! \ldots z_J!} \left(s^g_{1t}\right)^{z_1} \left(s^g_{2t}\right)^{z_2} \ldots \left(s^g_{Jt}\right)^{z_J}.$$

- The demand for each variety is a random variable

$$X^g_{jt} = \frac{z^g_{jt} \eta Y_t}{p^g_{jt} n_t}.$$

Back to our model: finite n_t

- Let z^g_{jt} be the number of purchases of variety j in good g.
- The vector $z^g_t = \{z^g_{jt} : j = 1, \ldots, J\}$ is distributed according to a Multinomial

$$\Pr(z) = \frac{n^g_t!}{z_1! z_2! \ldots z_J!} \left(s^g_{1t}\right)^{z_1} \left(s^g_{2t}\right)^{z_2} \ldots \left(s^g_{Jt}\right)^{z_J}.$$

- The demand for each variety is a random variable

$$X^g_{jt} = \frac{z^g_{jt} \eta Y_t}{p^g_{jt} n_t}.$$

- Conveniently, expenditures in good g are $Y^g_t = \alpha^g \eta Y_t$.

Not everything all the time

- Any trade flow can be zero with positive probability:

$$\Pr(z_{jt}^{g} = 0) = (1 - s_{jt}^{g})^{n_t^g}$$

Not everything all the time

- Any trade flow can be zero with positive probability:

$$\Pr(z_{jt}^g = 0) = (1 - s_{jt}^g)^{n_t^g}$$

- The universe of trade transactions in a year or five minutes still is a sample.
- The set of traded goods is no longer homogeneous along the frequency domain in our model.
  - Total trade value still scales proportionally with period length.

The extensive margin

- Let $d_{jt}^{g} = 0$ if $X_{jt}^{g} = 0$, $d_{jt}^{g} = 1$ otherwise.

The extensive margin

- Let $d_{jt}^g = 0$ if $X_{jt}^g = 0$, $d_{jt}^g = 1$ otherwise.
- The dummies d_{jt}^g are *not* independent.

The extensive margin

- Let $d^g_{jt} = 0$ if $X^g_{jt} = 0$, $d^g_{jt} = 1$ otherwise.
- The dummies d^g_{jt} are *not* independent.
- Expected number of traded varieties for good g at date t,

$$E\left\{\sum_J d^g_{jt}\right\} = J - \sum_J (1 - s^g_{jt})^{n^g_t}.$$

- or traded goods for country j,

$$E\left\{\sum_G d^g_{jt}\right\} = G - \sum_G (1 - s^g_{jt})^{n^g_t}.$$

The extensive margin - Dynamic

- Entry of variety j in good g at date $t+1$ is equivalent to the event $(1-d_{jt}^g)d_{jt+1}^g = 1$, thus

$$\Pr((1-d_{jt}^g)d_{jt+1}^g = 1) = (1-s_{jt}^g)^{n_t^g}\left(1-(1-s_{jt+1}^g)^{n_{t+1}^g}\right).$$

as purchases are independent across time.

The extensive margin - Dynamic

- Entry of variety j in good g at date $t+1$ is equivalent to the event $(1-d^g_{jt})d^g_{jt+1}=1$, thus

$$\Pr((1-d^g_{jt})d^g_{jt+1}=1)=(1-s^g_{jt})^{n^g_t}\left(1-(1-s^g_{jt+1})^{n^g_{t+1}}\right).$$

  as purchases are independent across time.

- Expected number of varieties j that enter in good g

$$E\left\{\sum_J(1-d^g_{jt})d^g_{jt+1}\right\}=\sum_J(1-s^g_{jt})^{n^g_t}\left(1-(1-s^g_{jt+1})^{n^g_{t+1}}\right).$$

- Similar formulas for exit and net change.

The extensive margin - Survival analysis

- The probability variety j in good g has positive sales for periods $t, t+1, \ldots, t+k$

$$\Pr\left(d_{jl}^{g} = 1 : l = t, \ldots, t+k\right) = \prod_{l=t}^{t+k}\left(1 - (1 - s_{jl}^{g})^{n_l^g}\right).$$

- Number of varieties expected to last at least k periods.

$$E\left\{\sum_{J}\prod_{l=t}^{t+k} d_{jl}^{g}\right\} = \sum_{J}\prod_{l=t}^{t+k}\left(1 - (1 - s_{jl}^{g})^{n_l^g}\right).$$

The extensive margin - Survival analysis

- The probability variety j in good g has positive sales for periods $t, t+1, \ldots, t+k$

$$\Pr\left(d_{jl}^{g} = 1 : l = t, \ldots, t+k\right) = \prod_{l=t}^{t+k}\left(1 - (1 - s_{jl}^{g})^{n_l^g}\right).$$

- Number of varieties expected to last at least k periods.

$$E\left\{\sum_{J}\prod_{l=t}^{t+k} d_{jl}^{g}\right\} = \sum_{J}\prod_{l=t}^{t+k}\left(1 - (1 - s_{jl}^{g})^{n_l^g}\right).$$

- Exit hazard rate is easily computed too but we need to resort to simulation to aggregate across goods.

The intensive margin

- Variation in z_{jt}^{g} drives the intensive margin.
  - Expenditures per purchase constant.
- Conditional on being traded, revenues are

$$E\left\{Y_{jt}^{g}|d_{jt}^{g}=1\right\}=\frac{s_{jt}^{g}}{1-(1-s_{jt}^{g})^{n_t^g}}Y_t^g.$$

- Easy to derive expectations conditional on a full set J^* of non-traded flows, and similar ...
- Useful to weigh entry and exit by value, characterize revenues per variety conditional on survival ...

Domestic varieties

- Domestic varieties: we do not observe them.

Domestic varieties

- Domestic varieties: we do not observe them.
- We treat the demand for domestic varieties as a unobserved parameter, n^g_{1t} — we do not need to.
- The model remains isomorphic with the remaining $J-1$ (foreign) varieties.
- Conditional on n^g_{1t}, the vector z^g_{-1t} is distributed according to a Multinomial with $n^g_t - n^g_{1t}$ draws and probabilities

$$\hat{s}^g_{jt} = \frac{s^g_{jt}}{1 - s^g_{1t}}.$$

Domestic varieties

- Domestic varieties: we do not observe them.
- We treat the demand for domestic varieties as a unobserved parameter, n_{1t}^g — we do not need to.
- The model remains isomorphic with the remaining $J-1$ (foreign) varieties.
- Conditional on n_{1t}^g, the vector z_{-1t}^g is distributed according to a Multinomial with $n_t^g - n_{1t}^g$ draws and probabilities

$$\hat{s}_{jt}^g = \frac{s_{jt}^g}{1-s_{1t}^g}.$$

- We thus focus on foreign varieties from now on.

Continuous time better?

- We have toyed with writing the model in continuous time.
- Purchase orders for good g arrive at a Poisson rate

$$\lambda_t^g = \alpha^g \lambda_t.$$

- This allows for a simpler, more elegant exposition and the model can be evaluated at any frequency.
- But we need to return to discrete time anyway for quantitative evaluation — much easier if the number of purchases is a parameter rather than a random variable.

Data and calibration

Data

- U.S. imports from 1990 to 2001.
- 10537 HS10 product codes after dropping:
  - New or obsolete codes in 1990-2001,
  - Special classification chapters 98 and 99,
  - Petroleum, fuels, and electricity.
- 120 countries accounting for more than 99 percent of total trade.
  - We append some gravity data from CEPII and World Bank.

Calibration

- We aim for a calibration as parsimonious as possible.

Calibration

- We aim for a calibration as parsimonious as possible.
- Set α^g to match the average expenditure share in good g over 1990-2001.

Calibration

- We aim for a calibration as parsimonious as possible.
- Set α^g to match the average expenditure share in good g over 1990-2001.
- We use the HS10-level elasticities of substitution estimated by Broda and Weinstein (2006).
    - Exploiting the link between σ^g and μ^g.
    - For missing elasticities, we impute the average elasticity at the 4-digit level.

Calibration - Prices

- We need to take a stand on all prices — including those for non-traded varieties.
- We posit

$$\ln p^g_{jt} = \ln \bar{p} + \ln \tau_j + \ln \tau_g + \ln \tau_t.$$

Calibration - Prices

- We need to take a stand on all prices — including those for non-traded varieties.
- We posit
$$\ln p_{jt}^{g} = \ln \bar{p} + \ln \tau_j + \ln \tau_g + \ln \tau_t.$$
- But only country fixed-effects matter. We set them to match the average share of products per country in the data.
  - Similar results if we use a gravity equation instead.
- The underlying assumption is the relative prices across countries are constant.
  - No comparative advantage forces.

Calibration - Purchase decisions

- The defining property of the purchase decisions is that they are *independent*.

Calibration - Purchase decisions

- The defining property of the purchase decisions is that they are *independent*.
- We set the initial level n_0 to match the share of traded commodities in 1990.

Calibration - Purchase decisions

- The defining property of the purchase decisions is that they are *independent*.
- We set the initial level n_0 to match the share of traded commodities in 1990.
- Growth rate of income γ_y set to overall U.S. GDP growth 1990-2000.
- Growth rate foreign purchases γ_n set to reproduce import growth 1990-2000.
    - Increased import penetration implies $\gamma_n > \gamma_y$.

Calibration - Parameter summary

Description	Parameter	Target/Source
Preferences across goods	$\{\alpha^g\}_G$	HS10 expenditure shares
Elasticities across goods	$\{\sigma^g\}_G$	Broda and Weinstein (2006)
Country fixed effects	$\{\tau_j\}_J$	Products traded per country
Purchases decisions in 1990	$n_0 = 16 \times 10^6$	Good-varieties traded in 1990
Growth rate income	$\gamma_y = .035$	Real GDP growth 1990-2001
Growth rate purchases	$\gamma_n = .065$	Import penetration 1990-2001
Preference non-tradeables	$\eta = .3$	Expenditure share non-tradeables

Table : **Summary calibration**

Cross-section results

Cross-section results

- Summarizing, the calibration matched:
  - The intensive margin across products,
  - The extensive margin across countries.

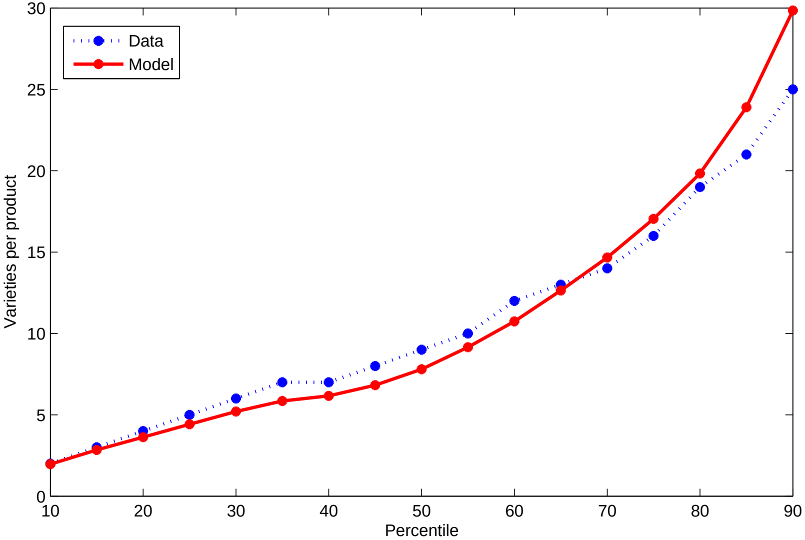
Cross-section results

- Summarizing, the calibration matched:
  - The intensive margin across products,
  - The extensive margin across countries.
- We start by checking the extensive margin across products.

Cross-section results

- Summarizing, the calibration matched:
  - The intensive margin across products,
  - The extensive margin across countries.
- We start by checking the extensive margin across products.
- We then present some further cuts,
  - Conditional on product size,
  - Conditional on product elasticity.

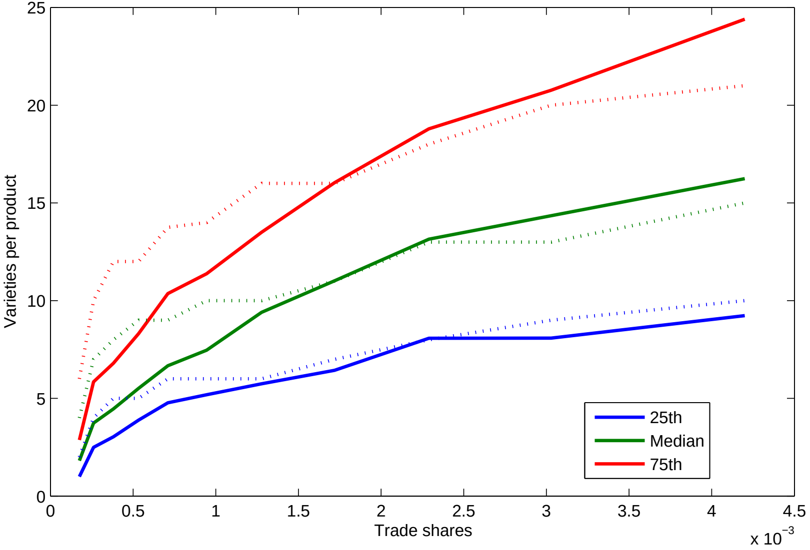
Distribution of number of varieties per product



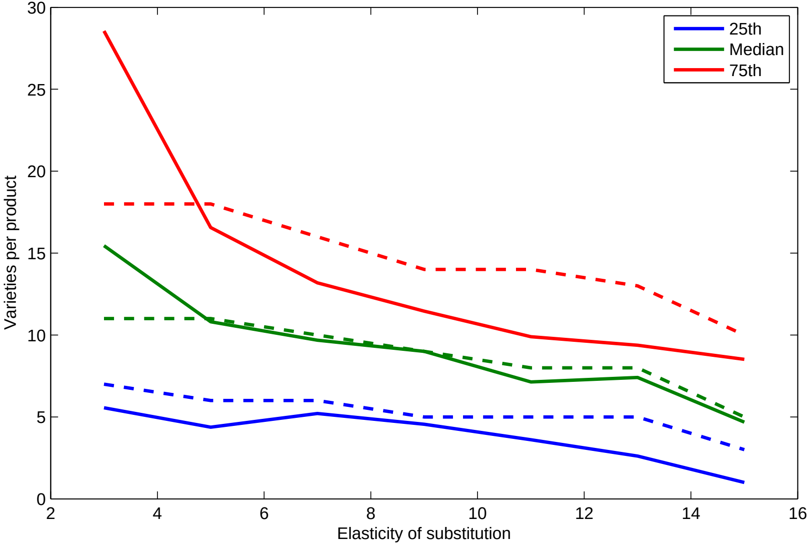
Do the right products have the right amount of varieties?

- In the model, the *expected* number of varieties across goods varies due to:
  - The market share of the good, α^g,
  - The elasticity, σ^g.
- In addition, the *realized* number of varieties is stochastic.
  - As long as we cluster enough products together, the sample variation washes down for all but the smallest bins.
- Next we present the distribution conditional on size and elasticity.

Conditional distribution - by market share



Conditional distribution - by elasticity



The extensive margin and the elasticity

- The relationship with the elasticity is intriguing.
  - There is no correlation between size and elasticity.

The extensive margin and the elasticity

- The relationship with the elasticity is intriguing.
  - There is no correlation between size and elasticity.
- In the standard CES demand, higher elasticity leads to a more skewed market share distribution.

The extensive margin and the elasticity

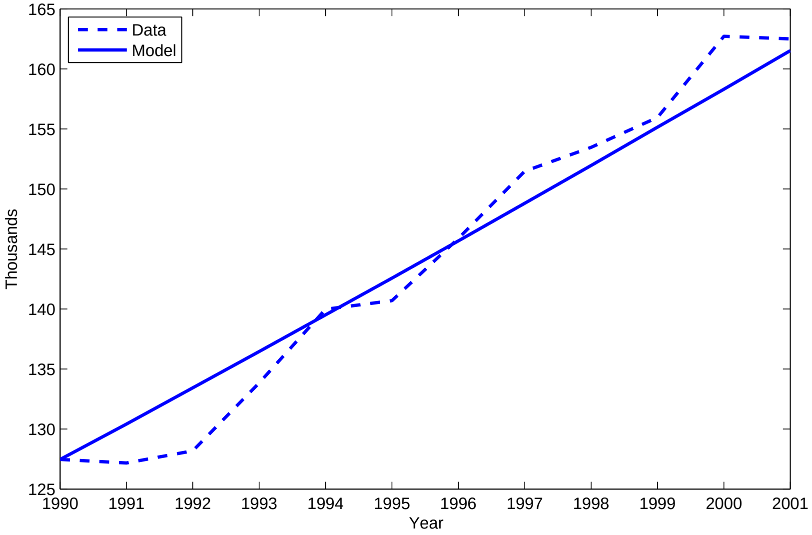
- The relationship with the elasticity is intriguing.
  - There is no correlation between size and elasticity.
- In the standard CES demand, higher elasticity leads to a more skewed market share distribution.
- In our model, it skews the underlying probability distribution and decreases the expected number of varieties.
  - Easier to get both heads and tails with a fair coin than with one with $.99$ chance of heads.
- Relationship is a little bit weak in the data — but elasticities are estimates, not point parameters as we have assumed.

Dynamic results

Questions across time

1. What is the contribution of the extensive margin to trade growth in the model and data?
2. Does the model reproduce the large amount of churning we observe in the data?
3. How do the survival probability and exit hazard rate in the model compare with the data?

Traded varieties over time



Extensive margin contribution

- Recall we targeted:
  - The initial (1990) level of traded varieties with n_0,
  - Increase in import penetration M_t/Y_t over 1990-2001 with the growth rate of purchases at $\gamma_n = .065$.
- We did **not** target the growth rate of varieties.

Extensive margin contribution

- Recall we targeted:
  - The initial (1990) level of traded varieties with n_0,
  - Increase in import penetration M_t/Y_t over 1990-2001 with the growth rate of purchases at $\gamma_n = .065$.
- We did **not** target the growth rate of varieties.
- How does the model get it spot on at $\gamma_d = .022$?

Extensive margin contribution - Skewness paramount

- If all varieties were symmetric, $s = 1/J$, we would have

$$\gamma_d \approx \gamma_n D/JG = .065 * .9 = .058,$$

overshooting the data.

Extensive margin contribution - Skewness paramount

- If all varieties were symmetric, $s = 1/J$, we would have

$$\gamma_d \approx \gamma_n D/JG = .065 * .9 = .058,$$

  overshooting the data.

- Crucially, the skewness in the underlying probability distribution "slows" down growth along the extensive margin.
  - New purchases being fulfilled by varieties previously traded count toward the intensive margin.
  - While most varieties are not traded, there are some that attract a lot of purchases.

Extensive margin contribution - Skewness paramount

- If all varieties were symmetric, $s = 1/J$, we would have

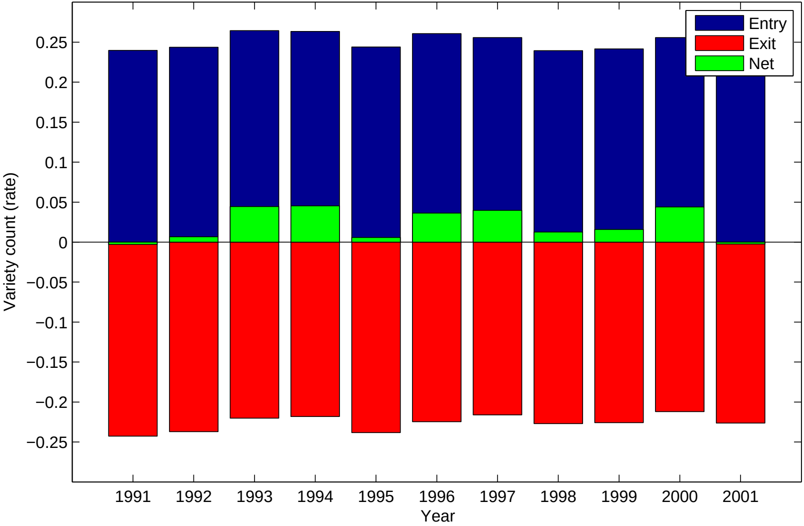
$$\gamma_d \approx \gamma_n D/JG = .065 * .9 = .058,$$

  overshooting the data.

- Crucially, the skewness in the underlying probability distribution "slows" down growth along the extensive margin.
  - New purchases being fulfilled by varieties previously traded count toward the intensive margin.
  - While most varieties are not traded, there are some that attract a lot of purchases.

  Getting γ_d requires getting the right amount of skewness.

A lot of churning of varieties



Entry and exit

	Data		Model	
	By count	By value	By count	By value
Entry	24.6 %	1.1 %	27.0 %	1.2 %
Exit	22.4 %	0.8 %	24.8 %	1.0 %
Net	2.2 %	0.3 %	2.2 %	0.3 %

Table : **Entry and exit**

The model's core

- We naturally reproduce two commanding observations in the data:
  - There are many zero trade flows,
  - There is a lot of churning.

The model's core

- We naturally reproduce two commanding observations in the data:
  - There are many zero trade flows,
  - There is a lot of churning.

- The model views actual trade in most country-product pairs as a *rare* event.
  1. Many zeros: most of time, nothing happens.
  2. Count data: z events more common than $z + 1$ ...
  3. High churning: continuous trade is twice as rare.

The model's core

- We naturally reproduce two commanding observations in the data:
  - There are many zero trade flows,
  - There is a lot of churning.
- The model views actual trade in most country-product pairs as a *rare* event.
  1. Many zeros: most of time, nothing happens.
  2. Count data: z events more common than $z + 1$ ...
  3. High churning: continuous trade is twice as rare.
- A good quantitative fit, though, requires more than this.
  - We have already mention the role of skewness in the cross-section and for net entry.

Churning, quantitatively

- Churning is low for a variety if either:
  - It is very likely to be traded, or
  - it is very unlikely to be traded.

Churning, quantitatively

- Churning is low for a variety if either:
  - It is very likely to be traded, or
  - it is very unlikely to be traded.
- Without growth, churning is given by $2p(1-p)$ where p is the probability of being traded.
  - Concave with p
  - Maximum at $p = .5$ or at full entropy.

Churning, quantitatively

- Churning is low for a variety if either:
  - It is very likely to be traded, or
  - it is very unlikely to be traded.
- Without growth, churning is given by $2p(1-p)$ where p is the probability of being traded.
  - Concave with p
  - Maximum at $p = .5$ or at full entropy.
- The model predictions on total churning depend on the *location* of the underlying probability distribution.

Churning in the right places

	Data		Model	
	Goods	Countries	Goods	Countries
Entry	.69	.91	.72	.93
Exit	.54	.91	.68	.93
Net	-.13	-.11	-.08	-.01

Table : **Correlations of entry and exit with good and country size**

- Rates defined over corresponding commodity space, J or G, by count.

Churning in the right places

- Why are exit and entry increasing in the market share of the good or country?

Churning in the right places

- Why are exit and entry increasing in the market share of the good or country?
- For most varieties, the likelihood of being traded is low—certainly below $.5$.
- Larger countries or goods increase the probability of a given commodity being traded ...

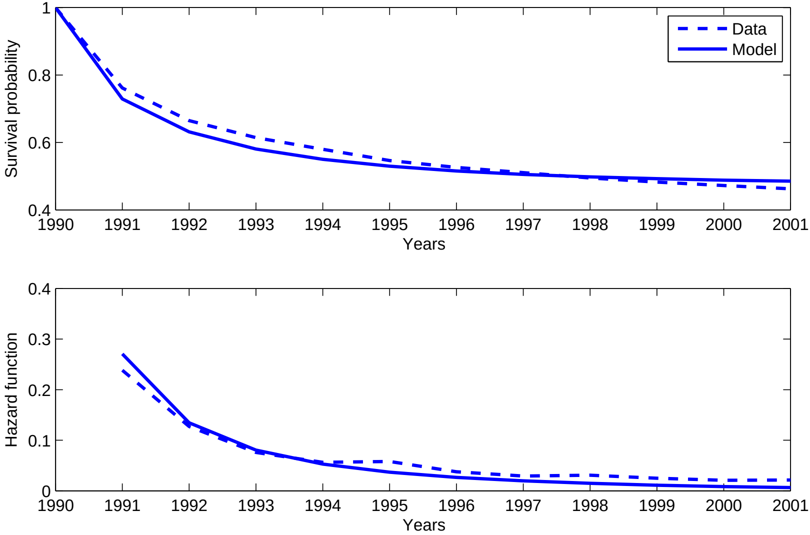
Churning in the right places

- Why are exit and entry increasing in the market share of the good or country?
- For most varieties, the likelihood of being traded is low—certainly below $.5$.
- Larger countries or goods increase the probability of a given commodity being traded ...
- ... which means the entropy is increasing, and so is the churning.
- The (weak) negative correlation with net entry is explained by the growth in purchases: the extensive margin contribution is smaller if more varieties are traded initially.

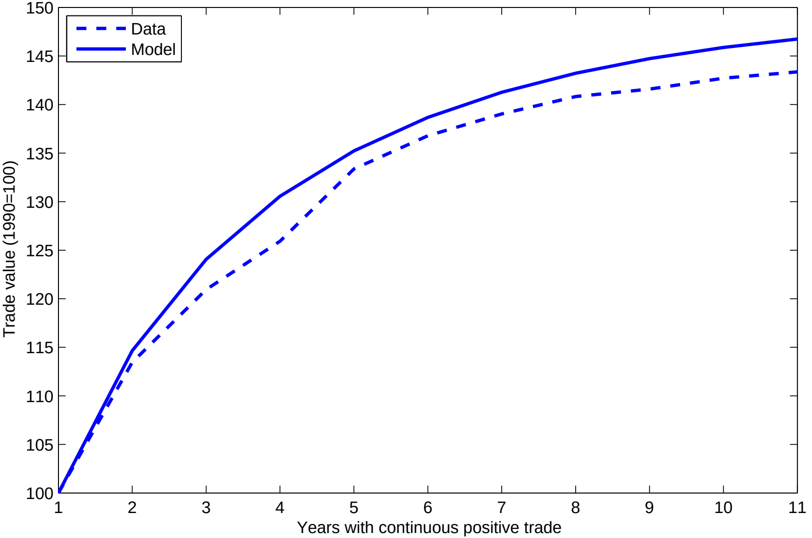
Survival analysis

- We now track the cohort of varieties traded in 1990 over time.
- What is the probability that a variety has been continuously traded (survival) since 1990?
- At what rate varieties stop being traded for the first time (exit hazard)?
- How does the composition of varieties change among survivors?

Survival probability and hazard rate



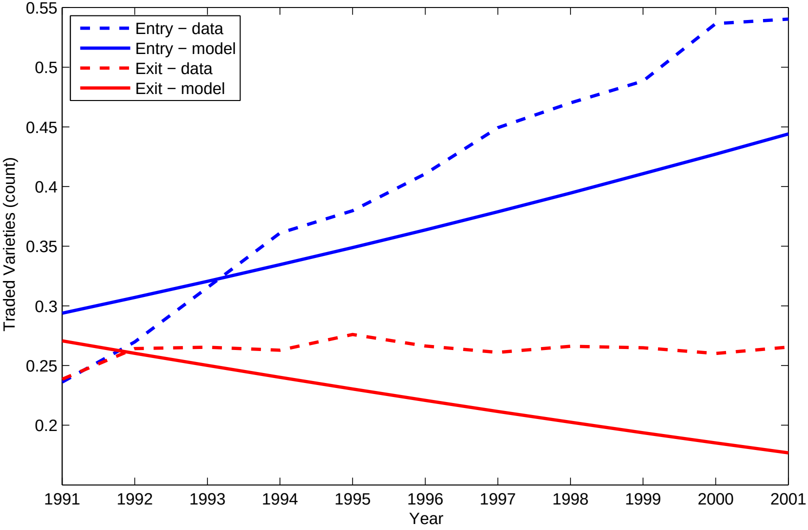
Trade value conditional on survival



Date-to-date entry and exit

- Now we compare the mix of traded varieties between a given year and 1990.
- If varieties re-enter or re-exit, they are counted.

Cumulative entry and exit since 1990



Date-to-date entry and exit

- Qualitatively the model is good, but we can no longer claim a quantitative fit.
- Partly, 1990-1991 is an outlier for net rates.
- However, the model seems to overstate the probability of re-entry substantially.
  - In the data, more varieties exit never to return.

Date-to-date entry and exit

- Qualitatively the model is good, but we can no longer claim a quantitative fit.
- Partly, 1990-1991 is an outlier for net rates.
- However, the model seems to overstate the probability of re-entry substantially.
  - In the data, more varieties exit never to return.
- This seems consistent with temporary price movements.
  - Mean-reversion in the underlying probability of being traded.
  - It may also be due to HS codes being re-defined.

The extensive margin during NAFTA

NAFTA

- NAFTA called for the phasing out, starting 1994, of virtually all restrictions on trade among the United States, Canada, and Mexico.
  - Romalis (2007), Kehoe and Ruhl (2012), Trefler (2004) ...
- Huge impact for Mexico. In just three years,
  - 70 % growth in U.S. imports,
  - 30 % more products, adding 10 % in value.
- Not so much for Canada.
  - Previous FTA (1989) already had reduced tariffs substantially.

NAFTA in our model

- We assumed relative prices stayed constant across countries over time.
  - Gravity is so strong that our assumption works pretty well.
- We will use NAFTA to see how the model fares with large, permanent relative price changes.

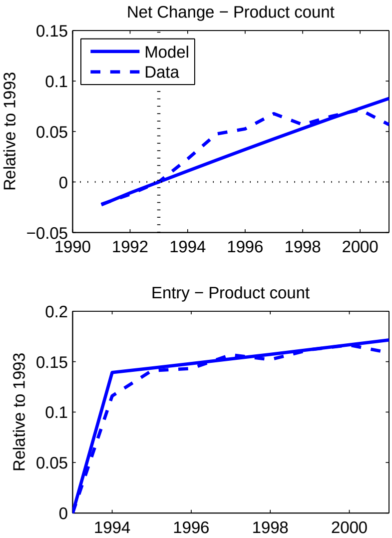
NAFTA in our model

- We assumed relative prices stayed constant across countries over time.
  - Gravity is so strong that our assumption works pretty well.
- We will use NAFTA to see how the model fares with large, permanent relative price changes.
- We first compare the baseline model (without NAFTA) with the data for Mexico and Canada, starting 1994.
  - We expect the model to miss for Mexico.

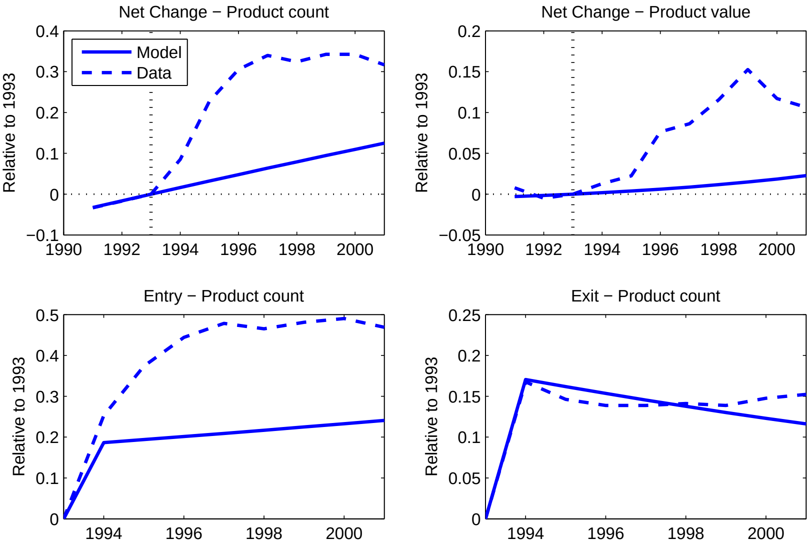
NAFTA in our model

- We assumed relative prices stayed constant across countries over time.
  - Gravity is so strong that our assumption works pretty well.
- We will use NAFTA to see how the model fares with large, permanent relative price changes.
- We first compare the baseline model (without NAFTA) with the data for Mexico and Canada, starting 1994.
  - We expect the model to miss for Mexico.
- We will then "add" NAFTA by adjusting Mexico's relative price and compare again model and data.

Canada after NAFTA: Model and Data



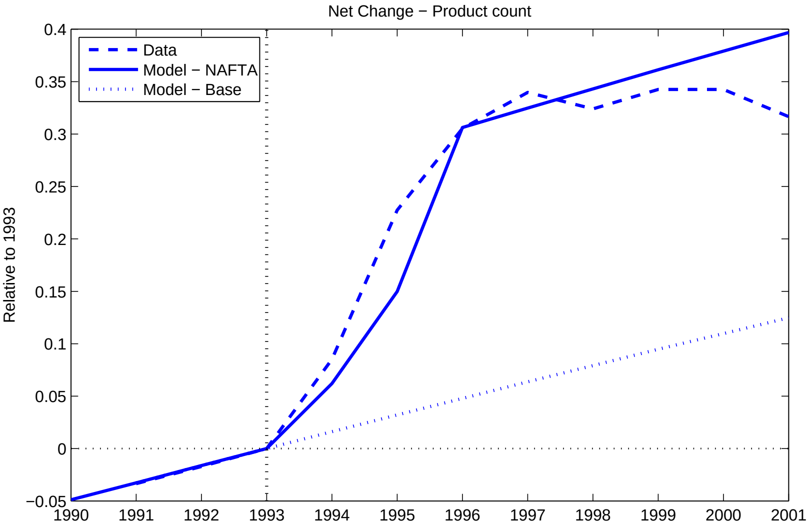
Mexico after NAFTA: Model and Data



Adding back NAFTA

- We cut Mexico's relative price by 10 percentage points.
  - In line with estimates elsewhere.
  - Matches well total trade increase (not surprising).
- We smooth the tariff cut over three years.
  - Half of tariffs were eliminated immediately in 1994,
  - Yet a substantial amount of goods were protected until 1999.
- Given that we match the total trade surge, do we get the extensive margin right?

Yes, we do



Net Change − Product count
Data
Model − NAFTA
Model − Base
Relative to 1993
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001

The extensive margin during NAFTA

- Our model gets the contribution of the extensive margin right naturally.
- As Mexican varieties get cheaper, they are purchased more *often*, so more of them are observed in a given year.

The extensive margin during NAFTA

- Our model gets the contribution of the extensive margin right naturally.
- As Mexican varieties get cheaper, they are purchased more *often*, so more of them are observed in a given year.
- Clearly some of the trade creation involved FDI, distribution networks, marketing . . .

The extensive margin during NAFTA

- Our model gets the contribution of the extensive margin right naturally.
- As Mexican varieties get cheaper, they are purchased more *often*, so more of them are observed in a given year.
- Clearly some of the trade creation involved FDI, distribution networks, marketing . . .
- It seems that trade was possible before, just not as frequent, without those investments.

Re-evaluating welfare gains from trade

New varieties and welfare gains

- Feenstra (1994) and Broda and Weinstein (2006) have attempted to measure the welfare gains due to newly-traded goods.
- The analysis is based on a CES demand system, and relies strongly on "everything all the time" to identify available commodities at any point in time.

New varieties and welfare gains

- Feenstra (1994) and Broda and Weinstein (2006) have attempted to measure the welfare gains due to newly-traded goods.
- The analysis is based on a CES demand system, and relies strongly on "everything all the time" to identify available commodities at any point in time.
- In our analysis all goods are available but not all traded.
  - Our model can be seen as the null hypothesis.

Feenstra price index

- Feenstra (1994) derives the CES welfare-based price index across two periods $t, t+1$, with changing varieties.
- Assume prices are constant across time: $p_{jt} = p_{jt-1}, \forall j$.
- Let I_t be the set of available varieties at date t,

$$I_t = \{i : X_{it} > 0\}$$

- Let I be the set of *common* varieties, $I = I_t \cap I_{t-1}$.
- Let λ_t be the share of *common* goods in date-t expenditures

$$\lambda_t = \frac{\sum_{j \in I} p_j X_{jt}}{\sum_{j \in I_t} p_j X_{jt}}.$$

Feenstra price index (2)

- The welfare-based price index is

$$\pi\left(X_{t-1}, X_t\right) = \left(\frac{\lambda_t}{\lambda_{t-1}}\right)^{\frac{1}{\sigma-1}}. \tag{1}$$

- Old goods lose in expenditure share $\lambda_t < \lambda_{t-1}$:
- New goods would have been purchased in large quantities *had they been available*.
- Weak axiom of revealed preference: consumer is happier.

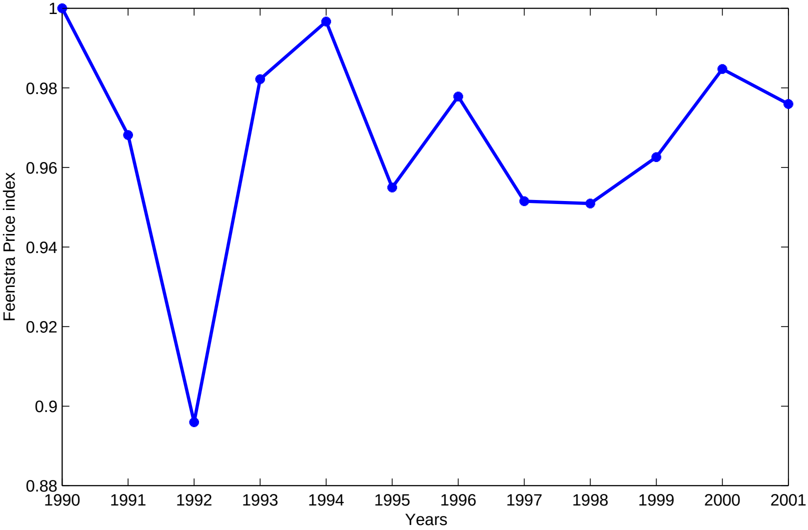
Feenstra price index under finite purchases

- With a finite number of purchases, π is a function of several random variables:
  - Number of purchases n_t,
  - Set of *traded* varieties, I_t,
  - Set of common varieties, I,
  - Expenditures X_{t-1}, X_t.
- Common set I implies that expenditure shares λ_{t-1}, λ_t are not independent even if X_{t-1}, X_t are.

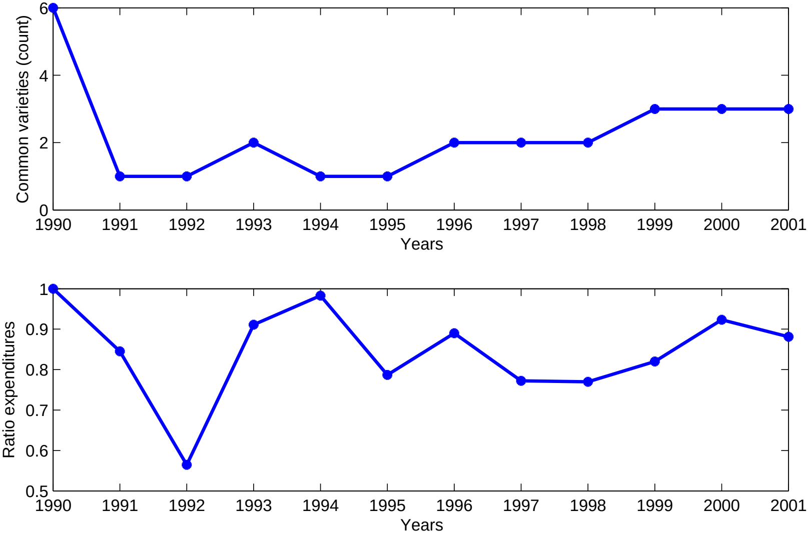
Feenstra price index under finite purchases

- With a finite number of purchases, π is a function of several random variables:
  - Number of purchases n_t,
  - Set of *traded* varieties, I_t,
  - Set of common varieties, I,
  - Expenditures X_{t-1}, X_t.
- Common set I implies that expenditure shares λ_{t-1}, λ_t are not independent even if X_{t-1}, X_t are.
- Some "prima facie" evidence on randomness.

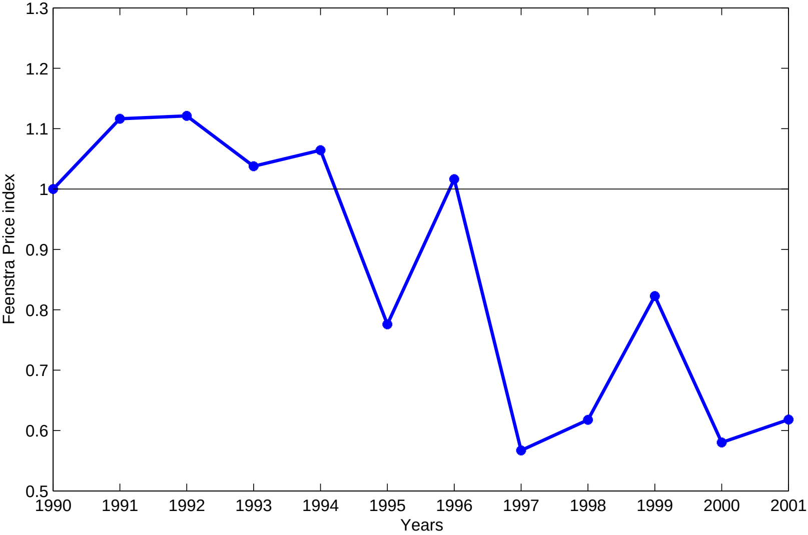
HS 1604194000: Fish sticks breaded/coated with batter or similarly prepared, not cooked, not in oil



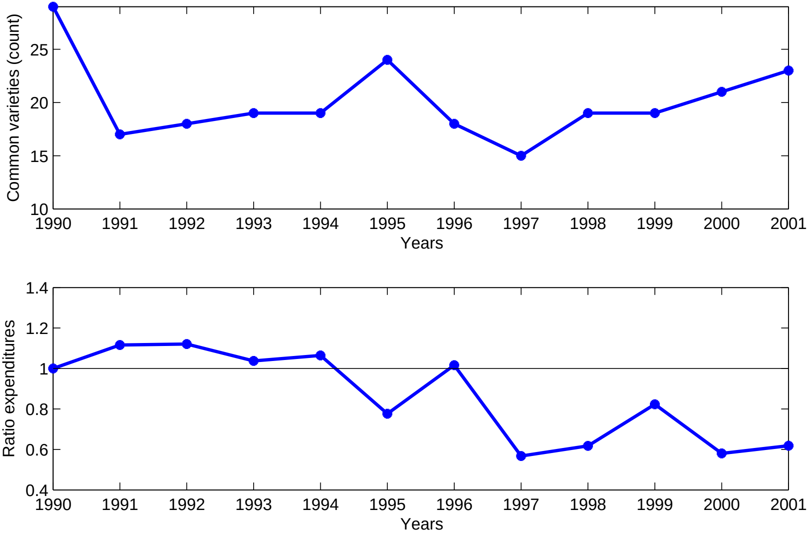
HS 1604194000: Fish sticks breaded/coated with batter or similarly prepared, not cooked, not in oil



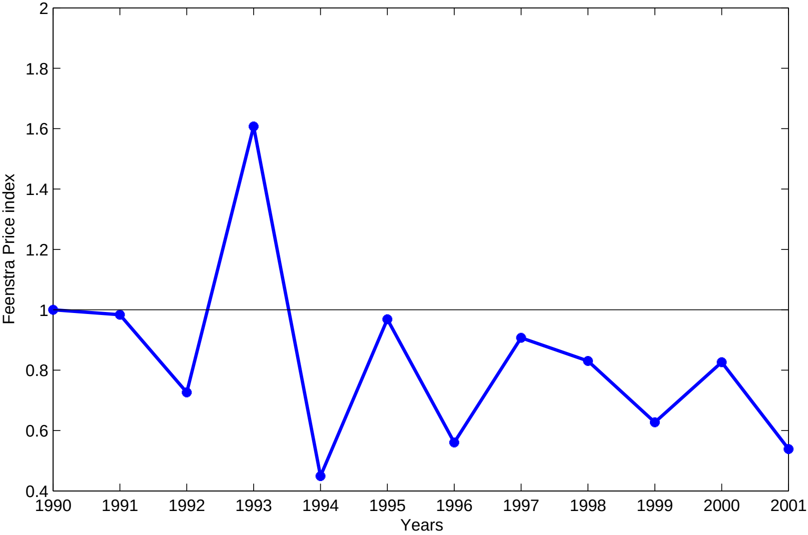
HS 6110102040: Girl's sweaters of other wool, knitted or crocheted



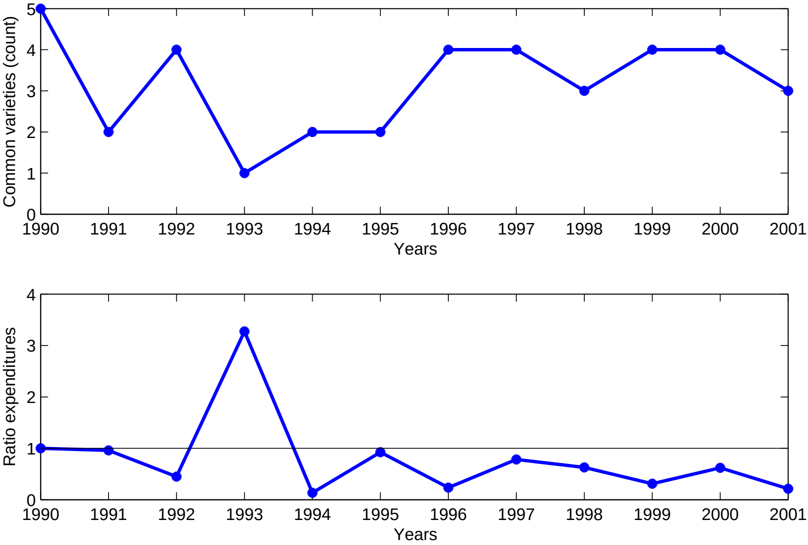
HS 6110102040: Girl's sweaters of other wool, knitted or crocheted



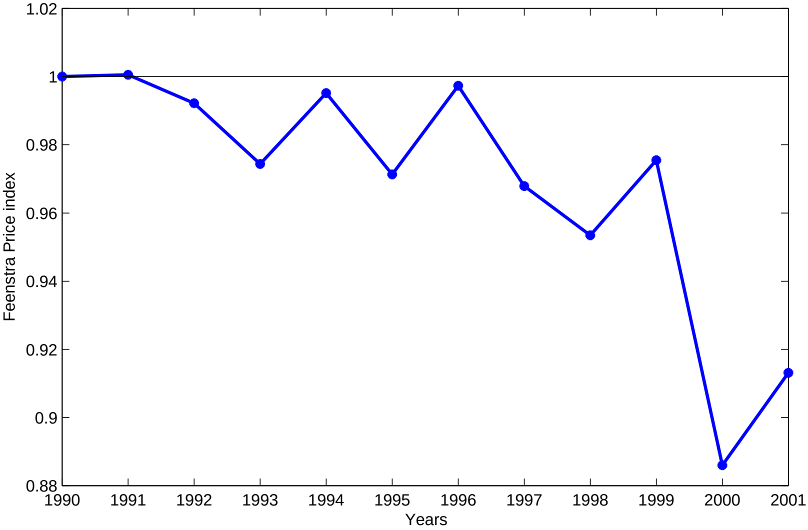
HS 6204433020: Girls dresses of synthetic fiber containing 36 % or more



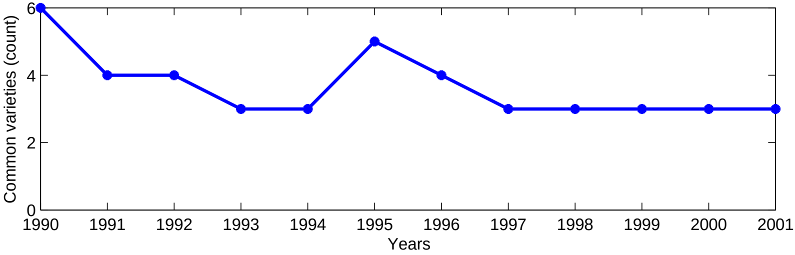
HS 6204433020: Girls dresses of synthetic fiber containing 36 % or more

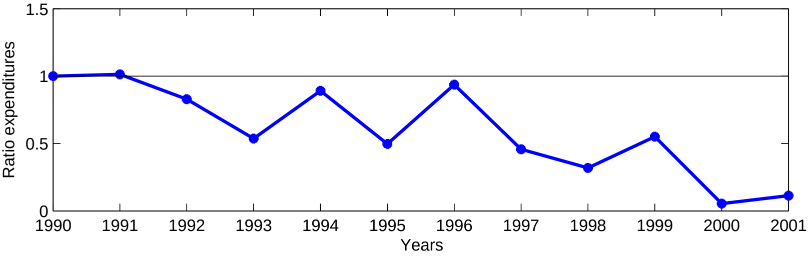


HS 9103108060: Battery for clocks, excluding travel clocks, NESOI



HS 9103108060: Battery for clocks, excluding travel clocks, NESOI





Upward Bias under $n_{t-1} = n_t$

- Take $n_{t-1} = n_t = n$ as given.
- Variables $\{I_{t-1}, I_t, \lambda_{t-1}, \lambda_t\}$ remain a random variable.

Upward Bias under $n_{t-1} = n_t$

- Take $n_{t-1} = n_t = n$ as given.
- Variables $\{I_{t-1}, I_t, \lambda_{t-1}, \lambda_t\}$ remain a random variable.
- Expenditure vectors X_t and X_{t-1} are i.i.d.
- Expenditure shares λ_t and λ_{t-1} are identically distributed but not independent.

Upward Bias under $n_{t-1} = n_t$

- Take $n_{t-1} = n_t = n$ as given.
- Variables $\{I_{t-1}, I_t, \lambda_{t-1}, \lambda_t\}$ remain a random variable.
- Expenditure vectors X_t and X_{t-1} are i.i.d.
- Expenditure shares λ_t and λ_{t-1} are identically distributed but not independent.
- Note that under H_0, all varieties are available in both periods. Yet the expenditure share is biased downwards:

$$E\{\lambda_t | n\} < 1.$$

Upward Bias under $n_{t-1} = n_t$

- Despite under-estimating the expenditure share in common varieties, the Feenstra price index is biased *upwards*:

$$E(\hat{\pi}|n) > 1$$

Upward Bias under $n_{t-1} = n_t$

- Despite under-estimating the expenditure share in common varieties, the Feenstra price index is biased *upwards*:

$$E(\hat{\pi}|n) > 1$$

- Proof: Note that

$$\hat{\pi}(X_{t-1}, X_t) = \frac{1}{\hat{\pi}(X_t, X_{t-1})}.$$

- And (X_{t-1}, X_t) and (X_t, X_{t-1}) are equally likely.
- Using Jensen's inequality,

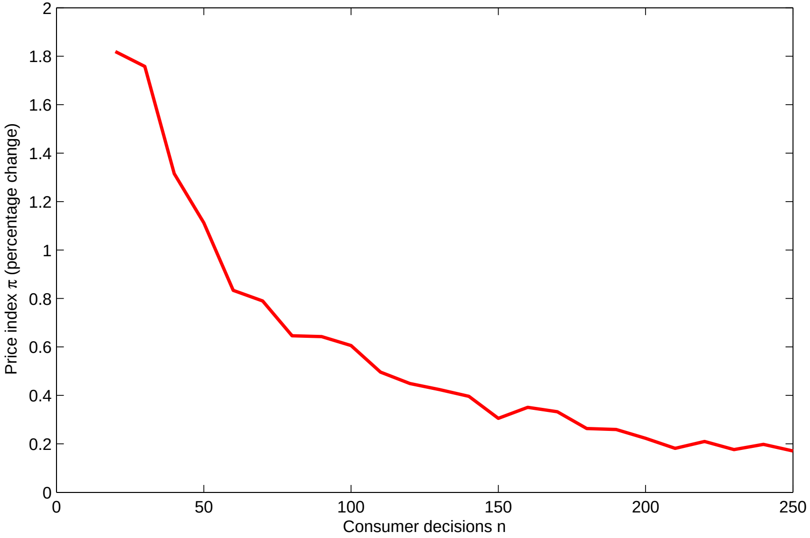
$$E(\hat{\pi}|n) = E\left(\frac{\hat{\pi} + 1/\hat{\pi}}{2}|n\right) > 1,$$

for any finite n.

How big is this bias?

- Quantitatively, the bias disappears very quickly with n.
  - Little convexity in $\pi + \pi^{-1}$ in the neighborhood of 1.
  - Sets I_{t-1}, I_t converge quite fast,
  - However λ_{t-1}, λ_t converges slowly to 1.
- Skewness in the underlying price distribution helps to keep the bias in check.
  - High-demand varieties rarely go unobserved, and make up a large share of total expenditures.
- Confidence intervals can remain quite wide.

Upward bias for $\hat{\pi}$ with $n_{t-1} = n_t = n$



Downward Bias under $n_{t-1} < n_t$

- Consider now growth in purchases, $n_{t-1} < n_t$
- Here the bias is *downwards*,

$$E(\hat{\pi}|n_{t-1}, n_t) < 1$$

  and can be sizeable if:
  - Growth rate is high,
  - Number of varieties is large,
  - Initial number of purchases is low.
- The bias does not vanish by averaging across products: *aggregate* welfare gains will be overstated.

Downward Bias under $n_{t-1} < n_t$ - Intuition

- Why does purchase growth bias the price index downwards?

Downward Bias under $n_{t-1} < n_t$ - Intuition

- Why does purchase growth bias the price index downwards?
- Fix outcome at date $t-1$ and let n_t grow.
- The set common varieties is contained in the set of goods consumed at $t-1$. Clearly

$$E\{z_{it-1}|z_{it-1} > 0\} > E\{z_{it-1}\}$$

- As n_t grows,

$$\lim_{n_t \to \infty} I = I_{t-1}.$$

- Thus the expenditure share of I at date $t-1$ tends to 1.

Downward Bias under $n_{t-1} < n_t$ - Intuition

- The expenditure share in date t will tend to

$$\lim_{n_t \to \infty} \lambda_t = \sum_{I_{t-1}} s_i.$$

Downward Bias under $n_{t-1} < n_t$ - Intuition

- The expenditure share in date t will tend to

$$\lim_{n_t \to \infty} \lambda_t = \sum_{I_{t-1}} s_i.$$

- The bias **does not go away** with n_t

$$\lim_{n_t \to \infty} E\left(\hat{\pi}\right) = \left(\sum_{I_{t-1}} s_i \right)^{\frac{1}{\sigma - 1}} < 1$$

Downward Bias under $n_{t-1} < n_t$ - Intuition

- The expenditure share in date t will tend to

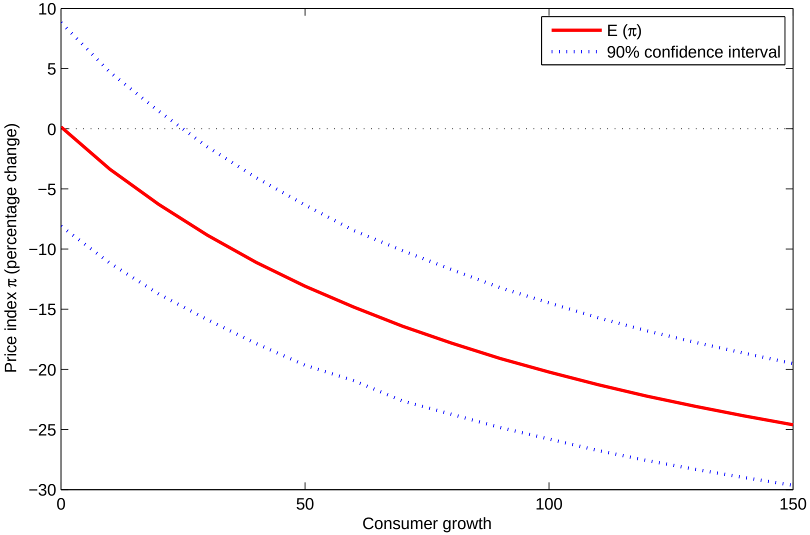
$$\lim_{n_t \to \infty} \lambda_t = \sum_{I_{t-1}} s_i.$$

- The bias **does not go away** with n_t

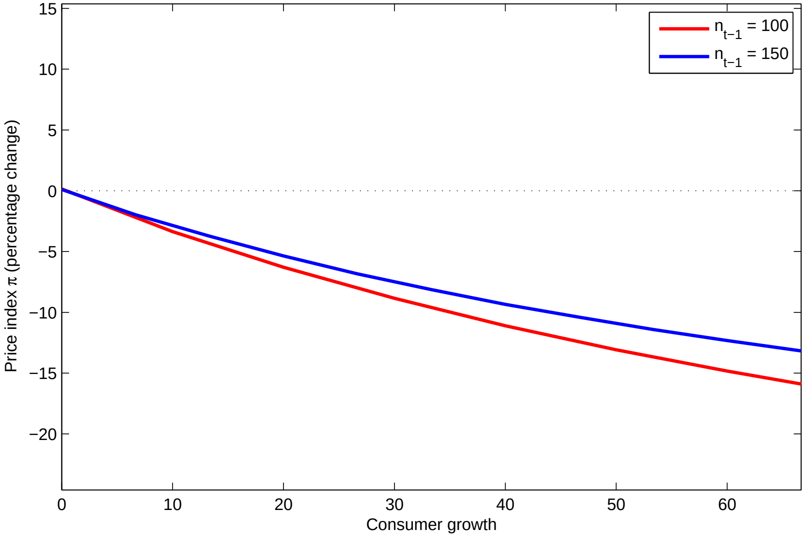
$$\lim_{n_t \to \infty} E\left(\hat{\pi}\right) = \left(\sum_{I_{t-1}} s_i\right)^{\frac{1}{\sigma-1}} < 1$$

- Moreover, the downward bias kicks in fast: if one new purchase chooses i with $X_{it} = 0, X_{it-1} > 0$:
  - λ_t increases by a single purchase,
  - λ_{t-1} increases by X_{it-1},
  - Price index $\hat{\pi}$ decreases.

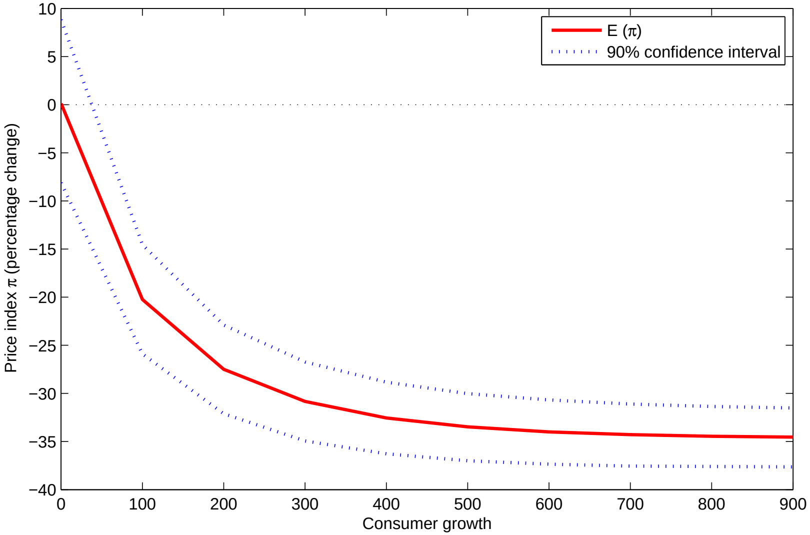
Downward bias for $\hat{\pi}$ with $n_{t-1} < n_t$



Downward bias for $\hat{\pi}$ with $n_{t-1} < n_t$ - Different n_{t-1}



Downward bias for $\hat{\pi}$ as $n_t \to \infty$



Feenstra (1994) import products

- Compare 8 product categories between 1967-1987.
- We match purchases in each date to the number of suppliers in each product.

	Suppliers		*Elasticity*
Import	1967	1987	σ
Men's Leather Athletic Shoes	16	27	6.2
Men's and boys' cotton knit shirts	24	51	5.8
Stainless steel bars	10	15	3.6
Carbon steel sheets	11	26	4.2
Color television receivers	6	15	8.4
Portable typewriters	17	14	3
Gold bullion*	19	32	27.2
Silver bullion	11	15	42.9

Feenstra (1994) import products - Results

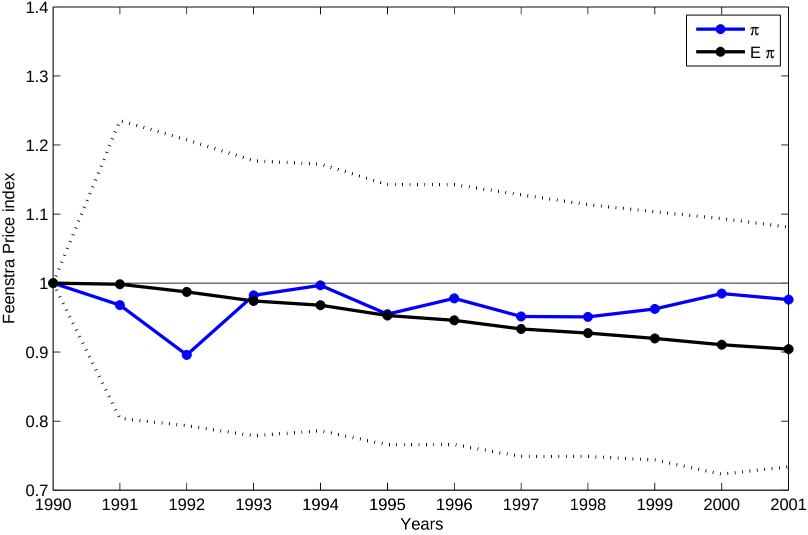
Import	Growth γ_n	*Cumulative Change in Import Price Index* Linear	$E\{\pi\}$	$\Pr(\pi < 1)$
Men's Leather Athletic Shoes	4.4 %	.9620	.9623	.9103
Men's and boys' cotton knit shirts	8.0 %	.9592	.9590	.9974
Stainless steel bars	3.1 %	.9241	.9293	.7601
Carbon steel sheets	7.2%	.89	.8896	.9678
Color television receivers	6.8 %	.9317	.982	.9231
Portable typewriters	-1.7%	1.0484	1.0549	.3277
Gold bullion[b]	9.5 %	.9931	.9931	.9389
Silver bullion	2.5 %	.9963	.9962	.7131

- Consistent downward bias across products:
  - Larger for low elasticity products,
  - Larger for fast growing products.
- We are possibly over-fitting n_{t-1}, n_t:
  - Actually reducing the sampling error.

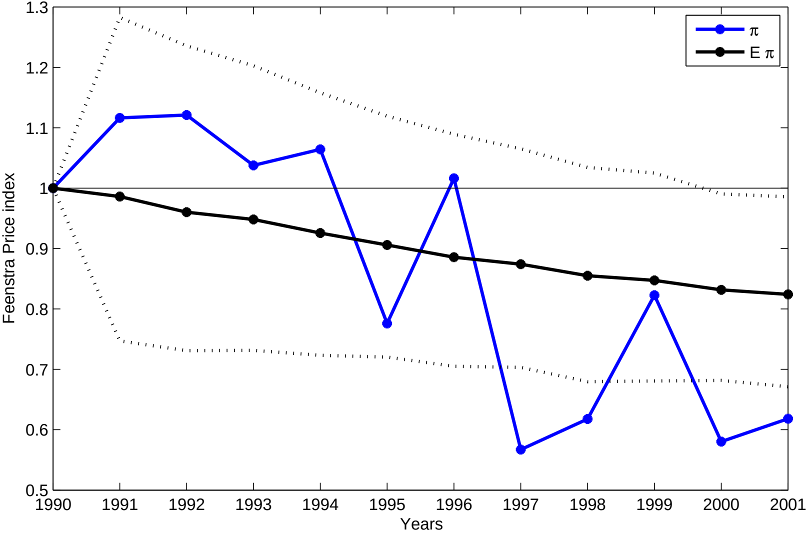
Testing variety gains in U.S. imports

- Broda and Weinstein (2006) measure the welfare gains from variety growth in U.S. imports.
  - Armington assumption: Each country provides a differentiated variety.
- We test whether the null hypothesis of no variety growth can be rejected for +10,000 HS10 product codes for the period 1990-2001.
- This is very much work in progress.

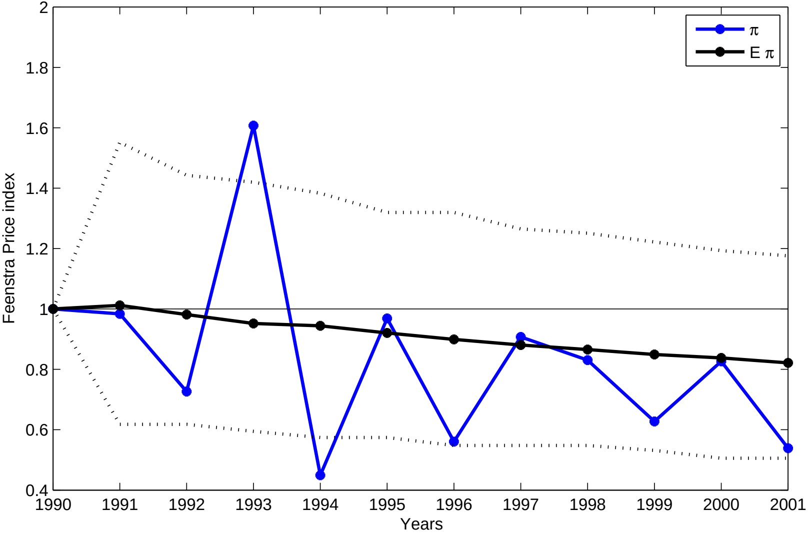
HS 1604194000: Fish sticks breaded/coated with batter or similarly prepared, not cooked, not in oil



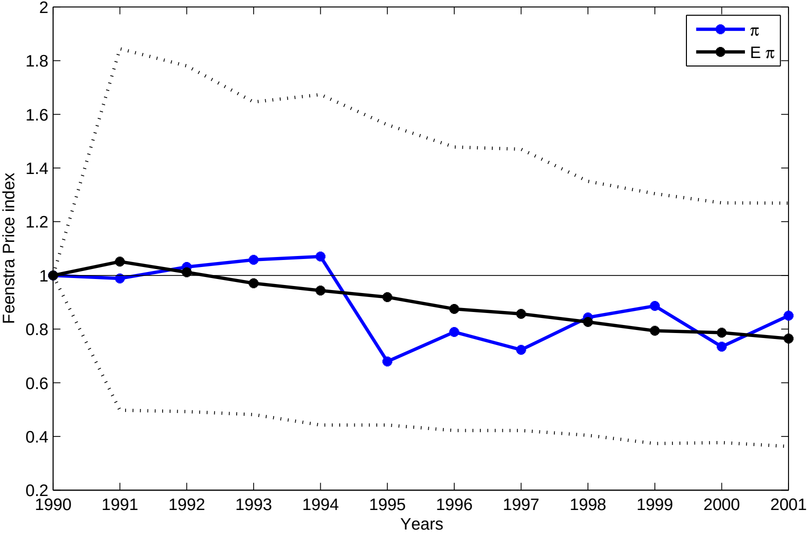
HS 6110102040: Girl's sweaters of other wool, knitted or crocheted



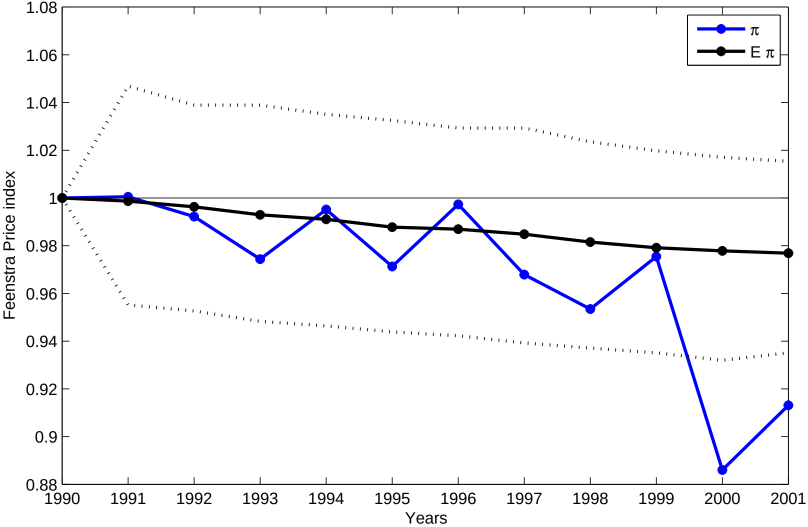
HS 6204433020: Girls dresses of synthetic fiber containing 36 % or more



HS 6204631510: Women's bib and brace overalls, synthetic fiber



HS 9103108060: Battery for clocks, excluding travel clocks, NESOI



Conclusions

Conclusions

- Our theory of the extensive margin is based on a simple discrete choice model on the demand side, evaluated for a finite number of purchases.
- Our calibration performs very well quantitatively
  - in the cross-section,
  - across time.
  - There are important implications for trade liberalization and welfare gains.

Conclusions

- Our theory of the extensive margin is based on a simple discrete choice model on the demand side, evaluated for a finite number of purchases.
- Our calibration performs very well quantitatively
  - in the cross-section,
  - across time.
  - There are important implications for trade liberalization and welfare gains.
- Economies of scale are conspicuously absent. Yet we believe them to matter at
  - Firm-level export participation,
  - Wholesale retail, with implications for total trade.