Misallocation and Product Choice

Stepan Gordeev
UConn

Sudhir Singh
MSU, World Bank

Jan 18, 2024

Texas Christian University

Misallocation and Product Choice

- **misallocation** is a big driver of cross-country income differences

Misallocation and Product Choice

- **misallocation** is a big driver of cross-country income differences
  - market frictions, distortionary policies $\rightarrow$ inputs not allocated to most productive uses

- **misallocation** is a big driver of cross-country income differences
  - market frictions, distortionary policies $\rightarrow$ inputs not allocated to most productive uses
- misallocation especially severe & costly in **agriculture** in low-income countries

Misallocation and Product Choice

- **misallocation** is a big driver of cross-country income differences
  - market frictions, distortionary policies $\rightarrow$ inputs not allocated to most productive uses
- misallocation especially severe & costly in **agriculture** in low-income countries
  - poor property rights, communal land distribution, hiring frictions, minimum prices

Misallocation and Product Choice

- **misallocation** is a big driver of cross-country income differences
  - market frictions, distortionary policies $\rightarrow$ inputs not allocated to most productive uses
- misallocation especially severe & costly in **agriculture** in low-income countries
  - poor property rights, communal land distribution, hiring frictions, minimum prices
- farms grow different products, most grow multiple

Misallocation and Product Choice

- **misallocation** is a big driver of cross-country income differences
  - market frictions, distortionary policies $\rightarrow$ inputs not allocated to most productive uses
- misallocation especially severe & costly in **agriculture** in low-income countries
  - poor property rights, communal land distribution, hiring frictions, minimum prices
- farms grow different products, most grow multiple
  - India: most common crop grown by **1/4** farms; **2/3** farms grow **2+** crops

Misallocation and Product Choice

- **misallocation** is a big driver of cross-country income differences
  - market frictions, distortionary policies $\rightarrow$ inputs not allocated to most productive uses
- misallocation especially severe & costly in **agriculture** in low-income countries
  - poor property rights, communal land distribution, hiring frictions, minimum prices
- farms grow different products, most grow multiple
  - India: most common crop grown by **1/4** farms; **2/3** farms grow **2+** crops
  - but misallocation literature uses **single-product** producers using the same production fn

Misallocation and Product Choice

- **misallocation** is a big driver of cross-country income differences
  - market frictions, distortionary policies $\rightarrow$ inputs not allocated to most productive uses
- misallocation especially severe & costly in **agriculture** in low-income countries
  - poor property rights, communal land distribution, hiring frictions, minimum prices
- farms grow different products, most grow multiple
  - India: most common crop grown by **1/4** farms; **2/3** farms grow **2+** crops
  - but misallocation literature uses **single-product** producers using the same production fn

how does heterogeneous product choice affect
aggregate misallocation cost?

Data

Preview

Data

- Indian farm-crop-level survey

DATA

- Indian farm-crop-level survey
- estimate product-level production functions
  - crops are significantly heterogeneous in input intensities, returns to scale

Preview

Data

- Indian farm-crop-level survey
- estimate product-level production functions
  - crops are significantly heterogeneous in input intensities, returns to scale

Model

Preview

Data

- Indian farm-crop-level survey
- estimate product-level production functions
  - crops are significantly heterogeneous in input intensities, returns to scale

Model

- multi-product farms choose products, face misallocative distortions

Preview

Data

- Indian farm-crop-level survey
- estimate product-level production functions
  - crops are significantly heterogeneous in input intensities, returns to scale

Model

- multi-product farms choose products, face misallocative distortions
- efficient input markets in India $\rightarrow$ aggregate agricultural output $\uparrow 4\times$

Preview

Data

- Indian farm-crop-level survey
- estimate product-level production functions
  - crops are significantly heterogeneous in input intensities, returns to scale

Model

- multi-product farms choose products, face misallocative distortions
- efficient input markets in India $\rightarrow$ aggregate agricultural output $\uparrow 4\times$
  - 1-product model: understate total cost of misallocation by **30%** ($3\times$ vs $4\times$ gain)

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

- misinterpret optimal crop heterogeneity as frictions $\rightarrow$ overstate misallocation

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

- misinterpret optimal crop heterogeneity as frictions $\rightarrow$ overstate misallocation
- miss endogenous product choice response to frictions $\rightarrow$ overstate misallocation

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

- misinterpret optimal crop heterogeneity as frictions $\rightarrow$ overstate misallocation
- miss endogenous product choice response to frictions $\rightarrow$ overstate misallocation
- miss within-farm productivity dispersion $\rightarrow$ understate misallocation

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

- misinterpret optimal crop heterogeneity as frictions $\rightarrow$ overstate misallocation
- miss endogenous product choice response to frictions $\rightarrow$ overstate misallocation
- miss within-farm productivity dispersion $\rightarrow$ understate misallocation
- ignore returns-to-scale heterogeneity between products $\rightarrow$ understate misallocation

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

- misinterpret optimal crop heterogeneity as frictions $\rightarrow$ overstate misallocation
- miss endogenous product choice response to frictions $\rightarrow$ overstate misallocation
- miss within-farm productivity dispersion $\rightarrow$ understate misallocation
- ignore returns-to-scale heterogeneity between products $\rightarrow$ understate misallocation

1-Product Model Error:

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

- misinterpret optimal crop heterogeneity as frictions → overstate misallocation
- miss endogenous product choice response to frictions → overstate misallocation
- miss within-farm productivity dispersion → understate misallocation
- ignore returns-to-scale heterogeneity between products → understate misallocation

1-Product Model Error:

- overstate frictions in data

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

- misinterpret optimal crop heterogeneity as frictions $\rightarrow$ overstate misallocation
- miss endogenous product choice response to frictions $\rightarrow$ overstate misallocation
- miss within-farm productivity dispersion $\rightarrow$ understate misallocation
- ignore returns-to-scale heterogeneity between products $\rightarrow$ understate misallocation

1-Product Model Error:

- overstate frictions in data
  - $\rightarrow$ overstate benefit of partial reallocations

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

- misinterpret optimal crop heterogeneity as frictions $\rightarrow$ overstate misallocation
- miss endogenous product choice response to frictions $\rightarrow$ overstate misallocation
- miss within-farm productivity dispersion $\rightarrow$ understate misallocation
- ignore returns-to-scale heterogeneity between products $\rightarrow$ understate misallocation

1-Product Model Error:

- overstate frictions in data
  - $\rightarrow$ overstate benefit of partial reallocations
- understate farm expansion if frictions lifted

Preview: Mechanisms

Compared to the Multi-Product Model, Standard 1-Product Models:

- misinterpret optimal crop heterogeneity as frictions → overstate misallocation
- miss endogenous product choice response to frictions → overstate misallocation
- miss within-farm productivity dispersion → understate misallocation
- ignore returns-to-scale heterogeneity between products → understate misallocation

1-Product Model Error:

- overstate frictions in data
  - → overstate benefit of partial reallocations
- understate farm expansion if frictions lifted
  - → understate total cost of misallocation

Literature

- Agricultural Misallocation

Literature

- **Agricultural Misallocation**
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing

Literature

- Agricultural Misallocation
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing
  - measure misallocation from specific frictions: *Chen (2017)*, *Gottlieb and Grobovsek (2019)*

Literature

- **Agricultural Misallocation**
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing
  - measure misallocation from specific frictions: *Chen (2017)*, *Gottlieb and Grobovsek (2019)*
  - identify overall misallocation from general frictions: *Gollin, Udry (2021)*, *Aragon, Restuccia, Rud (2022)*, *Adamopoulos, Brandt, Leight, Restuccia (2022)*, *Ayerst, Brandt, Restuccia (2023)*

LITERATURE

- AGRICULTURAL MISALLOCATION
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing
  - measure misallocation from specific frictions: *Chen (2017)*, *Gottlieb and Grobovsek (2019)*
  - identify overall misallocation from general frictions: *Gollin, Udry (2021)*, *Aragon, Restuccia, Rud (2022)*, *Adamopoulos, Brandt, Leight, Restuccia (2022)*, *Ayerst, Brandt, Restuccia (2023)*
  - this paper:

Literature

- **Agricultural Misallocation**
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing
  - measure misallocation from specific frictions: *Chen (2017)*, *Gottlieb and Grobovsek (2019)*
  - identify overall misallocation from general frictions: *Gollin, Udry (2021)*, *Aragon, Restuccia, Rud (2022)*, *Adamopoulos, Brandt, Leight, Restuccia (2022)*, *Ayerst, Brandt, Restuccia (2023)*
  - this paper:
    - > crop-specific production functions

LITERATURE

- AGRICULTURAL MISALLOCATION
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing
  - measure misallocation from specific frictions: *Chen (2017)*, *Gottlieb and Grobovsek (2019)*
  - identify overall misallocation from general frictions: *Gollin, Udry (2021)*, *Aragon, Restuccia, Rud (2022)*, *Adamopoulos, Brandt, Leight, Restuccia (2022)*, *Ayerst, Brandt, Restuccia (2023)*
  - this paper:
    - crop-specific production functions
    - multi-crop farms adjusting crop choice to frictions

Literature

- **Agricultural Misallocation**
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing
  - measure misallocation from specific frictions: *Chen (2017)*, *Gottlieb and Grobovsek (2019)*
  - identify overall misallocation from general frictions: *Gollin, Udry (2021)*, *Aragon, Restuccia, Rud (2022)*, *Adamopoulos, Brandt, Leight, Restuccia (2022)*, *Ayerst, Brandt, Restuccia (2023)*
  - this paper:
    - crop-specific production functions
    - multi-crop farms adjusting crop choice to frictions
- **Multi-Product Firms & Misallocation**

Literature

- **Agricultural Misallocation**
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing
  - measure misallocation from specific frictions: *Chen (2017)*, *Gottlieb and Grobovsek (2019)*
  - identify overall misallocation from general frictions: *Gollin, Udry (2021)*, *Aragon, Restuccia, Rud (2022)*, *Adamopoulos, Brandt, Leight, Restuccia (2022)*, *Ayerst, Brandt, Restuccia (2023)*
  - this paper:
    - > crop-specific production functions
    - > multi-crop farms adjusting crop choice to frictions
- **Multi-Product Firms & Misallocation**
  - *Jaef (2018)*, *Wang, Yang (2021)*: manufacturers choose # of products with heterog. TFPs
    - $\rightarrow$ # of products responds to distortions

Literature

- **Agricultural Misallocation**
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing
  - measure misallocation from specific frictions: *Chen (2017)*, *Gottlieb and Grobovsek (2019)*
  - identify overall misallocation from general frictions: *Gollin, Udry (2021)*, *Aragon, Restuccia, Rud (2022)*, *Adamopoulos, Brandt, Leight, Restuccia (2022)*, *Ayerst, Brandt, Restuccia (2023)*
  - this paper:
    - > crop-specific production functions
    - > multi-crop farms adjusting crop choice to frictions
- **Multi-Product Firms & Misallocation**
  - *Jaef (2018)*, *Wang, Yang (2021)*: manufacturers choose # of products with heterog. TFPs
    - → # of products responds to distortions
  - this paper:

LITERATURE

- **AGRICULTURAL MISALLOCATION**
  - base framework: *Chen, Restuccia, Santaeulalia-Llopis (2022)* building on *Hsieh, Klenow (2009)* for manufacturing
  - measure misallocation from specific frictions: *Chen (2017)*, *Gottlieb and Grobovsek (2019)*
  - identify overall misallocation from general frictions: *Gollin, Udry (2021)*, *Aragon, Restuccia, Rud (2022)*, *Adamopoulos, Brandt, Leight, Restuccia (2022)*, *Ayerst, Brandt, Restuccia (2023)*
  - this paper:
    - > crop-specific production functions
    - > multi-crop farms adjusting crop choice to frictions
- **MULTI-PRODUCT FIRMS & MISALLOCATION**
  - *Jaef (2018)*, *Wang, Yang (2021)*: manufacturers choose # of products with heterog. TFPs
    - → # of products responds to distortions
  - this paper:
    - > heterogeneous productivities and production functions

Why Agriculture?

- **agriculture** is the perfect setting to study **product choice and misallocation**

Why Agriculture?

- **agriculture** is the perfect setting to study **product choice and misallocation**
  - firm-product (farm-crop) inputs and outputs are feasible to measure

Why Agriculture?

- **agriculture** is the perfect setting to study **product choice and misallocation**
  - firm-product (farm-crop) inputs and outputs are feasible to measure
- but mechanisms apply to **non-agricultural** settings too
  - $\rightarrow$ relevant for *more* developed countries too

DATA

Data

- **India**'s Rural Economic and Demographic Survey (REDS)
  - nationally representative of rural Indian households

Data

- India's Rural Economic and Demographic Survey (REDS)
  - nationally representative of rural Indian households
  - use 2007-08 round, only with plot-crop-level data

Data

- India's Rural Economic and Demographic Survey (REDS)
  - nationally representative of rural Indian households
  - use 2007-08 round, only with plot-crop-level data
- 10,318 plots, 4,803 farmers

Data

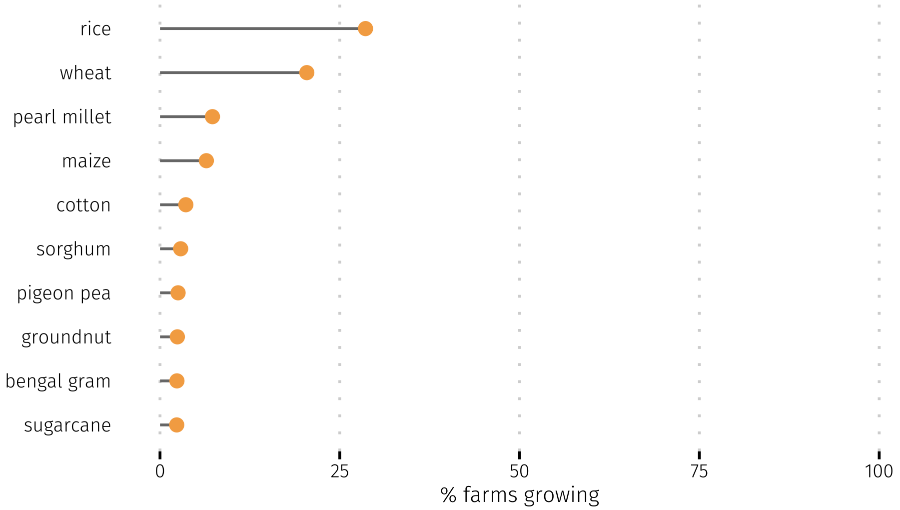
- India's Rural Economic and Demographic Survey (REDS)
  - nationally representative of rural Indian households
  - use 2007-08 round, only with plot-crop-level data
- 10,318 plots, 4,803 farmers
  - plot-crop-level inputs, outputs

Data

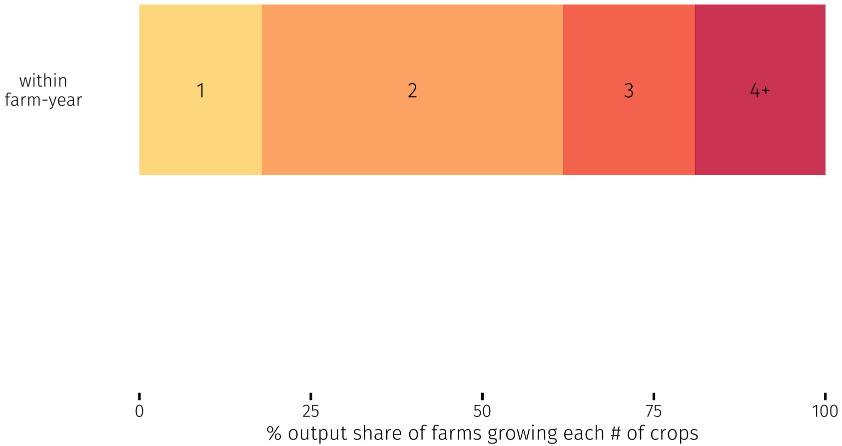
- India's Rural Economic and Demographic Survey (REDS)
  - nationally representative of rural Indian households
  - use 2007-08 round, only with plot-crop-level data
- 10,318 plots, 4,803 farmers
  - plot-crop-level inputs, outputs
  - plot-level physical characteristics

Heterogeneous Product Choice in India

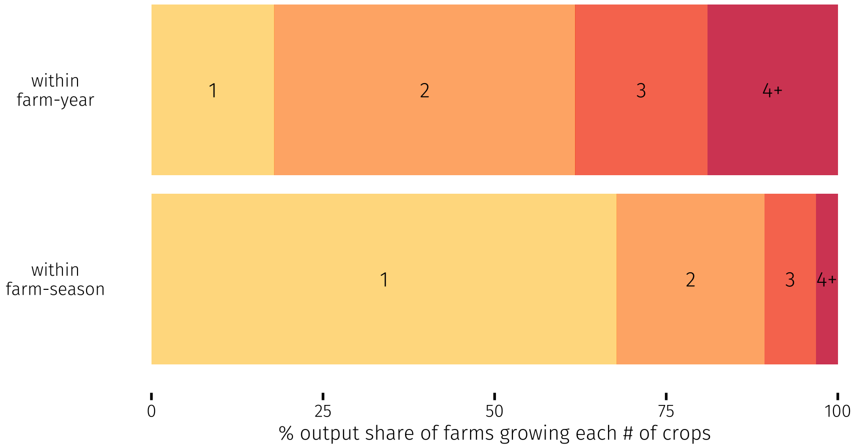
Crop Choice is Heterogeneous



Many Farms Grow Multiple Crops



Many Farms Grow Multiple Crops



- 3 agricultural seasons: Kharif (monsoon), Rabi (winter/spring), Zaid (summer/dry)

Production Functions

Estimate Crop-Specific Production Functions

- production function:

$$y_{f,i,t} = z_{f,i,t} l_{f,i,t}^{\gamma_i} x_{labor,f,i,t}^{\alpha_{labor,i}} x_{inter,f,i,t}^{\alpha_{inter,i}}$$

Estimate Crop-Specific Production Functions

- production function:

$$y_{f,i,t} = z_{f,i,t} l_{f,i,t}^{\gamma_i} x_{labor,f,i,t}^{\alpha_{labor,i}} x_{inter,f,i,t}^{\alpha_{inter,i}}$$

  - for farm f, crop i, season t

Estimate Crop-Specific Production Functions

- production function:

$$y_{f,i,t} = z_{f,i,t} l_{f,i,t}^{\gamma_i} x_{labor,f,i,t}^{\alpha_{labor,i}} x_{inter,f,i,t}^{\alpha_{inter,i}}$$

  - for farm f, crop i, season t
  - y = physical output

Estimate Crop-Specific Production Functions

- production function:

$$y_{f,i,t} = z_{f,i,t} l_{f,i,t}^{\gamma_i} x_{labor,f,i,t}^{\alpha_{labor,i}} x_{inter,f,i,t}^{\alpha_{inter,i}}$$

  - for farm f, crop i, season t
  - y = physical output
  - l = land input (quality-adjusted)

Estimate Crop-Specific Production Functions

- production function:

$$y_{f,i,t} = z_{f,i,t} l_{f,i,t}^{\gamma_i} x_{labor,f,i,t}^{\alpha_{labor,i}} x_{inter,f,i,t}^{\alpha_{inter,i}}$$

  - for farm f, crop i, season t
  - y = physical output
  - l = land input (quality-adjusted)
  - x_{labor} = labor input (days)

Estimate Crop-Specific Production Functions

- production function:

$$y_{f,i,t} = z_{f,i,t} l_{f,i,t}^{\gamma_i} x_{labor,f,i,t}^{\alpha_{labor,i}} x_{inter,f,i,t}^{\alpha_{inter,i}}$$

  - for farm f, crop i, season t
  - y = physical output
  - l = land input (quality-adjusted)
  - x_{labor} = labor input (days)
  - x_{inter} = intermediate inputs (seeds, fertilizer, etc.)
    - *inputs, output*

Estimate Crop-Specific Production Functions

- production function:

$$y_{f,i,t} = z_{f,i,t} l_{f,i,t}^{\gamma_i} x_{labor,f,i,t}^{\alpha_{labor,i}} x_{inter,f,i,t}^{\alpha_{inter,i}}$$

  - for farm *f*, crop *i*, season *t*
  - y = physical output
  - l = land input (quality-adjusted)
  - x_{labor} = labor input (days)
  - x_{inter} = intermediate inputs (seeds, fertilizer, etc.)
    - ▶ *inputs, output*
- merge crops into **5 groups**:
  - rice
  - wheat
  - other cereals
  - pulses
  - vegetables, fruits, oilseeds
    - ▶ *list of crops*

Inputs: Land

- common land input measures are problematic

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...
  - irrigation access: presence of wells, canals, ...

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...
  - irrigation access: presence of wells, canals, ...
- use random forests for prediction ▶ *details*

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...
  - irrigation access: presence of wells, canals, ...
- use random forests for prediction ▶ *details*
  - supervised machine learning algorithm

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...
  - irrigation access: presence of wells, canals, ...
- use random forests for prediction ▶ *details*
  - supervised machine learning algorithm
  - naturally adapts to non-linearities and interactions in the data

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...
  - irrigation access: presence of wells, canals, ...
- use random forests for prediction ▶ *details*
  - supervised machine learning algorithm
  - naturally adapts to non-linearities and interactions in the data
  - excels in out-of-sample prediction

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...
  - irrigation access: presence of wells, canals, ...
- use random forests for prediction ▶ *details*
  - supervised machine learning algorithm
  - naturally adapts to non-linearities and interactions in the data
  - excels in out-of-sample prediction
- $\rightarrow$ quality index = predicted price/acre

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...
  - irrigation access: presence of wells, canals, ...
- use random forests for prediction ▶ *details*
  - supervised machine learning algorithm
  - naturally adapts to non-linearities and interactions in the data
  - excels in out-of-sample prediction
- $\rightarrow$ quality index $=$ predicted price/acre
  - **captures quality**: RF explains $> \frac{1}{2}$ of observed price variation in **test** sample

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...
  - irrigation access: presence of wells, canals, ...
- use random forests for prediction ▶ *details*
  - supervised machine learning algorithm
  - naturally adapts to non-linearities and interactions in the data
  - excels in out-of-sample prediction
- $\rightarrow$ quality index = predicted price/acre
  - **captures quality**: RF explains $> \frac{1}{2}$ of observed price variation in **test** sample
  - **removes noise**: use predicted values, minimize overfitting $\rightarrow$ captured variation driven by population patterns, not individual mismeasurement

Inputs: Land

- common land input measures are problematic
  - **plot area**: ignores quality
  - **plot market price**: likely excessively noisy when land markets are underdeveloped
- instead: predict plot's market price with its physical land features
  - reg **log price/acre** on **plot features**
  - physical characteristics: soil type, color, salinity, drainage, ...
  - irrigation access: presence of wells, canals, ...
- use random forests for prediction ▶ *details*
  - supervised machine learning algorithm
  - naturally adapts to non-linearities and interactions in the data
  - excels in out-of-sample prediction
- $\rightarrow$ quality index = predicted price/acre
  - **captures quality**: RF explains $> \frac{1}{2}$ of observed price variation in **test** sample
  - **removes noise**: use predicted values, minimize overfitting $\rightarrow$ captured variation driven by population patterns, not individual mismeasurement
- quality-adjusted land = quality index $\times$ plot area

Instrumental Variables

- specification:

$$\log y_{j,i,t} = \gamma_i \log l^{\gamma_i}_{j,i,t} + \alpha_{labor,i} \log x_{labor,j,i,t} + \alpha_{inter,i} \log x_{inter,j,i,t} + \epsilon_{j,i,t}$$

  - for plot j, crop i, season t

- specification:

$$\log y_{j,i,t} = \gamma_i \log l^{\gamma_i}_{j,i,t} + \alpha_{labor,i} \log x_{labor,j,i,t} + \alpha_{inter,i} \log x_{inter,j,i,t} + \epsilon_{j,i,t}$$

  - for plot j, crop i, season t
- **production function simultaneity bias**: inputs are correlated with productivity

Instrumental Variables

- specification:

$$\log y_{j,i,t} = \gamma_i \log l^{\gamma_i}_{j,i,t} + \alpha_{labor,i} \log x_{labor,j,i,t} + \alpha_{inter,i} \log x_{inter,j,i,t} + \epsilon_{j,i,t}$$

  – for plot j, crop i, season t
- **production function simultaneity bias**: inputs are correlated with productivity
- **solution**: 2SLS using instruments for inputs adapting *Gollin, Udry (2021)*'s method

Instrumental Variables

- specification:

$$\log y_{j,i,t} = \gamma_i \log l_{j,i,t}^{\gamma_i} + \alpha_{labor,i} \log x_{labor,j,i,t} + \alpha_{inter,i} \log x_{inter,j,i,t} + \epsilon_{j,i,t}$$

  - for plot j, crop i, season t
- **production function simultaneity bias**: inputs are correlated with productivity
- **solution**: 2SLS using instruments for inputs adapting *Gollin, Udry (2021)*'s method
- **idea**: shocks to farm f's plots $k \neq j$ change shadow price of inputs on f's plot j

- specification:

$$\log y_{j,i,t} = \gamma_i \log l_{j,i,t}^{\gamma_i} + \alpha_{labor,i} \log x_{labor,j,i,t} + \alpha_{inter,i} \log x_{inter,j,i,t} + \epsilon_{j,i,t}$$

  - for plot j, crop i, season t
- **production function simultaneity bias**: inputs are correlated with productivity
- **solution**: 2SLS using instruments for inputs adapting *Gollin, Udry (2021)*'s method
- **idea**: shocks to farm f's plots $k \neq j$ change shadow price of inputs on f's plot j

  $\rightarrow$ serve as instruments for inputs on j

Instrumental Variables

- specification:

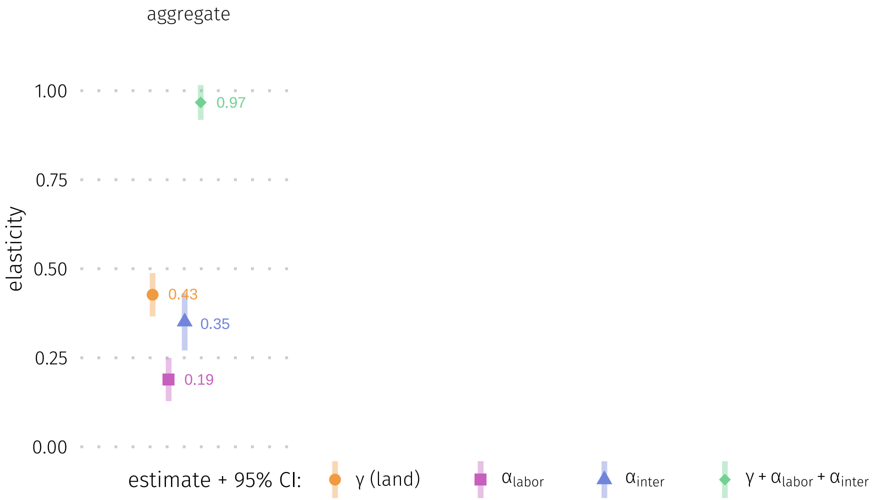
$$\log y_{j,i,t} = \gamma_i \log l^{\gamma_i}_{j,i,t} + \alpha_{labor,i} \log x_{labor,j,i,t} + \alpha_{inter,i} \log x_{inter,j,i,t} + \epsilon_{j,i,t}$$

  – for plot *j*, crop *i*, season *t*
- **production function simultaneity bias**: inputs are correlated with productivity
- **solution**: 2SLS using instruments for inputs adapting *Gollin, Udry (2021)*'s method
- **idea**: shocks to farm *f*'s plots $k \neq j$ change shadow price of inputs on *f*'s plot *j*

  → serve as instruments for inputs on *j*

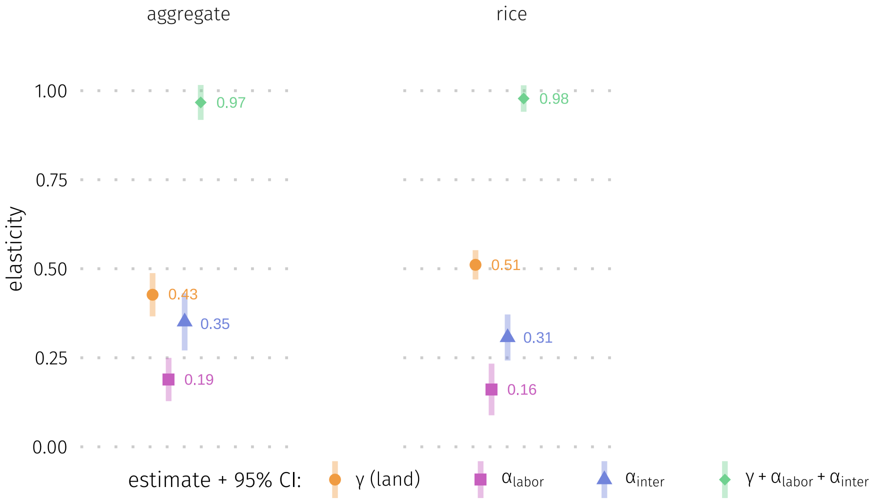
▶ *details*

Production Functions are Heterogeneous across Crops



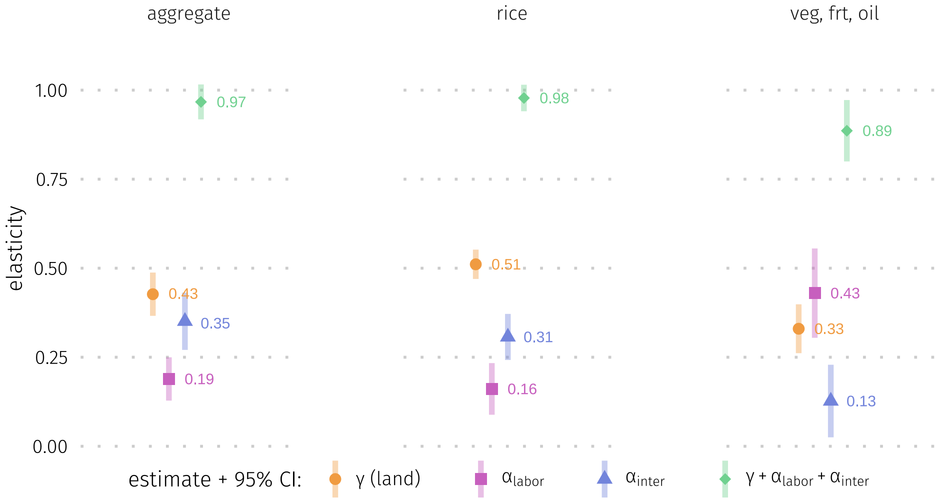
▶ *all crops* ▶ *table* ▶ *equality tests*

Production Functions are Heterogeneous across Crops



▶ *all crops* ▶ *table* ▶ *equality tests*

Production Functions are Heterogeneous across Crops



▶ *all crops* ▶ *table* ▶ *equality tests*

Model

Model Objectives

- OBJECTIVES:

Model Objectives

- **OBJECTIVES:**
  - model multi-product farm decisions in presence of distortions

Model Objectives

- **OBJECTIVES:**
  - model multi-product farm decisions in presence of distortions
  - provide a mapping from observable outcomes to unobserved distortions

Model Objectives

- **OBJECTIVES:**
  - model multi-product farm decisions in presence of distortions
  - provide a mapping from observable outcomes to unobserved distortions
  - quantify the aggregate output cost of misallocation induced by distortions

Model Objectives

- **OBJECTIVES:**
  - model multi-product farm decisions in presence of distortions
  - provide a mapping from observable outcomes to unobserved distortions
  - quantify the aggregate output cost of misallocation induced by distortions
- build on models of single-product firm-level misallocation

Model Objectives

- OBJECTIVES:
  - model multi-product farm decisions in presence of distortions
  - provide a mapping from observable outcomes to unobserved distortions
  - quantify the aggregate output cost of misallocation induced by distortions
- build on models of single-product firm-level misallocation
  - *Hsieh, Klenow (2009)*: misallocation in **manufacturing**

Model Objectives

- **OBJECTIVES:**
  - model multi-product farm decisions in presence of distortions
  - provide a mapping from observable outcomes to unobserved distortions
  - quantify the aggregate output cost of misallocation induced by distortions
- build on models of single-product firm-level misallocation
  - *Hsieh, Klenow (2009)*: misallocation in **manufacturing**
  - *Chen, Restuccia, Santaeulalia-Llopis (2022)*: misallocation in **agriculture**

Model Objectives

- **OBJECTIVES:**
  - model multi-product farm decisions in presence of distortions
  - provide a mapping from observable outcomes to unobserved distortions
  - quantify the aggregate output cost of misallocation induced by distortions
- build on models of single-product firm-level misallocation
  - *Hsieh, Klenow (2009)*: misallocation in **manufacturing**
  - *Chen, Restuccia, Santaeulalia-Llopis (2022)*: misallocation in **agriculture**
- add multi-product farms choosing among heterogeneous products

Farm: Production

$$\max \quad \underbrace{p \; \overbrace{z_f \; l_f{}^{\gamma} \, \Pi_g \left(x_{f,g}{}^{\alpha_g} \right)}^{y_f}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \; x_{f,g}}_{\text{flex. input costs}}$$

$$\text{s.t.} \qquad l_f = L_f \qquad (\lambda_f)$$

- profit-maximizing farm f: sells output py_f, pays for inputs

Farm: Production

$$\max \quad \underbrace{p \, z_f \, \overbrace{l_f^{\ \gamma} \Pi_g \left(x_{f,g}^{\ \alpha_g}\right)}^{y_f}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \, x_{f,g}}_{\text{flex. input costs}}$$

$$\text{s.t.} \quad l_f = L_f \quad (\lambda_f)$$

- profit-maximizing farm f: sells output py_f, pays for inputs
- Cobb-Douglas production function with TFP z_f

Farm: Production

$$\max \quad \underbrace{p \overbrace{z_f \, l_f^{\ \gamma} \, \Pi_g \left(x_{f,g}^{\ \ \alpha_g}\right)}^{y_f}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \qquad x_{f,g}}_{\text{flex. input costs}}$$

$$\text{s.t.} \qquad l_f \qquad = L_f \qquad (\lambda_f)$$

- profit-maximizing farm f: sells output py_f, pays for inputs
- Cobb-Douglas production function with TFP z_f
- flexible inputs g: labor, intermediates
  - quantity $x_{f,g}$ rented at r_g

Farm: Production

$$\max \quad \underbrace{p\,\overbrace{z_f\, l_f^{\ \gamma}\, \Pi_g \left(x_{f,g}^{\ \alpha_g}\right)}^{y_f}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \qquad x_{f,g}}_{\text{flex. input costs}}$$

$$\text{s.t.} \qquad l_f \qquad = L_f \qquad (\lambda_f)$$

- profit-maximizing farm f: sells output py_f, pays for inputs
- Cobb-Douglas production function with TFP z_f
- flexible inputs g: labor, intermediates
  - quantity $x_{f,g}$ rented at r_g
- land input l is in fixed supply L_f
  - almost no land market in India

Farm: Production

$$\max \quad \underbrace{p \, \overbrace{z_f \; l_f^{\;\gamma} \, \Pi_g \left(x_{f,g}^{\;\alpha_g}\right)}^{y_f}}_{\text{revenue}} \; - \; \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \quad x_{f,g}}_{\text{flex. input costs}}$$

$$\text{s.t.} \qquad l_f \qquad = L_f \qquad (\lambda_f)$$

- farm-input distortions $\tau_{f,g}$ capture misallocative frictions
  - represented with tax ($\tau_{f,g} > 1$) or subsidy ($\tau_{f,g} < 1$) idiosyncratic to farm f, input g

Farm: Distortions

$$\max \quad \underbrace{p \, \overbrace{z_f \, l_f^{\ \gamma} \, \Pi_g \left(x_{f,g}^{\ \alpha_g}\right)}^{y_f}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \, x_{f,g}}_{\text{flex. input costs}}$$

$$\text{s.t.} \quad l_f = L_f \quad (\lambda_f)$$

- farm-input distortions $\tau_{f,g}$ capture misallocative frictions
  - represented with tax ($\tau_{f,g} > 1$) or subsidy ($\tau_{f,g} < 1$) idiosyncratic to farm f, input g
- **between inputs:** $\tau_{f,labor} > \tau_{f,inter} \rightarrow$ farm f uses "too little" labor
  - e.g. hiring frictions, credit constraints, adjustment costs, transport costs

Farm: Distortions

$$\max \quad \underbrace{p \, \overbrace{z_f \; l_f^{\;\gamma} \, \Pi_g \left(x_{f,g}^{\;\alpha_g}\right)}^{y_f}}_{\text{revenue}} \; - \; \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \qquad x_{f,g}}_{\text{flex. input costs}}$$

$$\text{s.t.} \qquad l_f \qquad = L_f \qquad (\lambda_f)$$

- farm-input distortions $\tau_{f,g}$ capture misallocative frictions
  - represented with tax ($\tau_{f,g} > 1$) or subsidy ($\tau_{f,g} < 1$) idiosyncratic to farm f, input g
- **between inputs:** $\tau_{f,labor} > \tau_{f,inter}$ $\rightarrow$ farm f uses "too little" labor
  - e.g. hiring frictions, credit constraints, adjustment costs, transport costs
- **between farms:** $\tau_{1,g} > \tau_{2,g} \; \forall g$ $\rightarrow$ farm 1 is "too small" given its TFP z_f
  - e.g. market power, gov't subsidies, corruption

Farm: Distortions

$$\max \quad \underbrace{p\, \overbrace{z_f\, l_f^{\ \gamma}\, \Pi_g \left(x_{f,g}^{\ \alpha_g}\right)}^{y_f}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g}\, x_{f,g}}_{\text{flex. input costs}}$$

$$\text{s.t.} \quad l_f = L_f \quad (\lambda_f)$$

- farm-input distortions $\tau_{f,g}$ capture misallocative frictions
  - represented with tax ($\tau_{f,g} > 1$) or subsidy ($\tau_{f,g} < 1$) idiosyncratic to farm f, input g
- **between inputs:** $\tau_{f,labor} > \tau_{f,inter}$ $\rightarrow$ farm f uses "too little" labor
  - e.g. hiring frictions, credit constraints, adjustment costs, transport costs
- **between farms:** $\tau_{1,g} > \tau_{2,g}\ \forall g$ $\rightarrow$ farm 1 is "too small" given its TFP z_f
  - e.g. market power, gov't subsidies, corruption
- L_f fixed $\rightarrow$ land is also distorted unless distributed to equalize λ_f
  - e.g. lacking property rights, communal land distribution

Farm: Multiple Products

$$\max \underbrace{\sum_{i=1}^{N} p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- heterogeneous crops $i = 1 \dots N$

Farm: Multiple Products

$$\max \underbrace{\sum_{i=1}^{N} p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}} \right)}^{y_{f,i}}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- heterogeneous crops $i = 1 \ldots N$
- l in fixed supply $L_f \rightarrow$ interdependent crop production
  - params of crop i change $\rightarrow \lambda_f$ changes $\rightarrow$ inputs and outputs of crops $-i$ change
  - *Just, Zilberman, and Hochman (1983)*, *Shumway, Pope, Nash (1984)*
  - justifies *Gollin, Udry (2021)* prod. fn. identification

Farm: Multiple Products

$$\max \underbrace{\sum_{i=1}^{N} p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- heterogeneous crops $i = 1 \ldots N$
- l in fixed supply $L_f \rightarrow$ interdependent crop production
  - params of crop i change $\rightarrow \lambda_f$ changes $\rightarrow$ inputs and outputs of crops $-i$ change
  - *Just, Zilberman, and Hochman (1983)*, *Shumway, Pope, Nash (1984)*
  - justifies *Gollin, Udry (2021)* prod. fn. identification
- farm-input-crop distortions $\tau_{f,g,i}$, $\tau_{f,l,i}$

  $\rightarrow$ fit observed input ratio heterogeneity across crops within a farm

Farm: Fixed Cost

$$\max \underbrace{\sum_{i=1}^{N} p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}} - \underbrace{\sum_{i=1}^{N} \omega \cdot 1[y_{f,i} > 0]}_{\text{fixed cost per crop}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- fixed cost ω per produced crop
  - → farms choose **crop set** in addition to **crop mix**
  - – farms don't all produce everything
  - → fit observed heterogeneity in crop sets

Farm: Concavity

$$\max \underbrace{\sum_{i=1}^{N}\Big(p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}\Big)^{\eta}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}} - \underbrace{\sum_{i=1}^{N} \omega \cdot 1[y_{f,i} > 0]}_{\text{fixed cost per crop}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- crop-level concavity term $\eta < 1$

Farm: Concavity

$$\max \underbrace{\sum_{i=1}^{N}\Big(p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}\Big)^{\eta}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}} - \underbrace{\sum_{i=1}^{N} \omega \cdot 1[y_{f,i} > 0]}_{\text{fixed cost per crop}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- crop-level concavity term $\eta < 1$
- captures several motives for farmers to mix crops beyond prod. fn. DRS

Farm: Concavity

$$\max \underbrace{\sum_{i=1}^{N}\Big(p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}\Big)^{\eta}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}} - \underbrace{\sum_{i=1}^{N} \omega \cdot 1[y_{f,i} > 0]}_{\text{fixed cost per crop}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- crop-level concavity term $\eta < 1$
- captures several motives for farmers to mix crops beyond prod. fn. DRS
  - risk

Farm: Concavity

$$\max \underbrace{\sum_{i=1}^{N}\left(p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}\right)^{\eta}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}} - \underbrace{\sum_{i=1}^{N} \omega \cdot 1[y_{f,i} > 0]}_{\text{fixed cost per crop}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- crop-level concavity term $\eta < 1$
- captures several motives for farmers to mix crops beyond prod. fn. DRS
  - risk
  - subsistence + love of variety

Farm: Concavity

$$\max \underbrace{\sum_{i=1}^{N}\Big(p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}\Big)^{\eta}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}} - \underbrace{\sum_{i=1}^{N} \omega \cdot 1[y_{f,i} > 0]}_{\text{fixed cost per crop}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- crop-level concavity term $\eta < 1$
- captures several motives for farmers to mix crops beyond prod. fn. DRS
  - risk
  - subsistence + love of variety
  - market power

Farm: Concavity

$$\max \underbrace{\sum_{i=1}^{N}\left(p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}\right)^{\eta}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}} - \underbrace{\sum_{i=1}^{N} \omega \cdot 1[y_{f,i} > 0]}_{\text{fixed cost per crop}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- crop-level concavity term $\eta < 1$
- captures several motives for farmers to mix crops beyond prod. fn. DRS
  - risk
  - subsistence + love of variety
  - market power
  - farm's problem with η equivalent to subsistence or market power ▶ *details*

Farm: Concavity

$$\max \underbrace{\sum_{i=1}^{N}\left(p_i \overbrace{z_{f,i} {l_{f,i}}^{\gamma_i} \Pi_g \left({x_{f,g,i}}^{\alpha_{g,i}}\right)}^{y_{f,i}}\right)^{\eta}}_{\text{revenue}} - \underbrace{\sum_{g=1}^{G} r_g \tau_{f,g} \sum_{i=1}^{N} \tau_{f,g,i} x_{f,g,i}}_{\text{flex. input costs}} - \underbrace{\sum_{i=1}^{N} \omega \cdot 1[y_{f,i} > 0]}_{\text{fixed cost per crop}}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} l_{f,i} \tau_{f,l,i} = L_f \qquad (\lambda_f)$$

- crop-level concavity term $\eta < 1$
- captures several motives for farmers to mix crops beyond prod. fn. DRS
  - risk
  - subsistence + love of variety
  - market power
  - farm's problem with η equivalent to subsistence or market power ▶ *details*

▶ *solution*

- representative consumer buys crops, sells inputs, receives profit from owned farms

$$\max_{\{C_i\}_{i=1}^N} \left(\sum_i \varphi_i C_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

s.t.

$$\sum_i p_i C_i = \sum_g r_g X_g^{agg} + \Pi$$

- representative consumer buys crops, sells inputs, receives profit from owned farms

$$\max_{\{C_i\}_{i=1}^{N}} \left(\sum_i \varphi_i C_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

s.t.

$$\sum_i p_i C_i = \sum_g r_g X_g^{agg} + \Pi$$

- goods markets clear $\forall i$:

$$C_i = \sum_f y_{f,i}$$

- representative consumer buys crops, sells inputs, receives profit from owned farms

$$\max_{\{C_i\}_{i=1}^N} \left(\sum_i \varphi_i C_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

s.t.

$$\sum_i p_i C_i = \sum_g r_g X_g^{agg} + \Pi$$

- goods markets clear $\forall i$:

$$C_i = \sum_f y_{f,i}$$

- inputs markets clear $\forall g$:

$$\sum_f \sum_i x_{f,g,i} = X_g^{agg}$$

General Equilibrium

- representative consumer buys crops, sells inputs, receives profit from owned farms

$$\max_{\{C_i\}_{i=1}^N} \left(\sum_i \varphi_i C_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

s.t.

$$\sum_i p_i C_i = \sum_g r_g X_g^{agg} + \Pi$$

- goods markets clear $\forall i$:

$$C_i = \sum_f y_{f,i}$$

- inputs markets clear $\forall g$:

$$\sum_f \sum_i x_{f,g,i} = X_g^{agg}$$

- estimate $\sigma = 1.7$ from consumption FOC ▶ *details*

- representative consumer buys crops, sells inputs, receives profit from owned farms

$$\max_{\{C_i\}_{i=1}^N} \left(\sum_i \varphi_i C_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

s.t.

$$\sum_i p_i C_i = \sum_g r_g X_g^{agg} + \Pi$$

- goods markets clear $\forall i$:

$$C_i = \sum_f y_{f,i}$$

- inputs markets clear $\forall g$:

$$\sum_f \sum_i x_{f,g,i} = X_g^{agg}$$

- estimate $\sigma = 1.7$ from consumption FOC ▶ *details*

▶ *profits details*

Extracting Distortions

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

Extracting Distortions

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

- physical **productivity** implied by production fn.:

$$z_{f,i} = \frac{y_{f,i}}{l_{f,i}^{\gamma_i} \Pi_g \left(x_{f,g,i}^{\alpha_{g,i}}\right)}$$

Extracting Distortions

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

- physical **productivity** implied by production fn.:

$$z_{f,i} = \frac{y_{f,i}}{l_{f,i}^{\gamma_i} \Pi_g \left(x_{f,g,i}^{\alpha_{g,i}}\right)}$$

- $\rightarrow$ extracted fundamentals rationalize observed dispersion b/w farms

Extracting Distortions

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

- physical **productivity** implied by production fn.:

$$z_{f,i} = \frac{y_{f,i}}{l_{f,i}^{\gamma_i} \Pi_g \left(x_{f,g,i}^{\alpha_{g,i}}\right)}$$

- $\rightarrow$ extracted fundamentals rationalize observed dispersion b/w farms frictionless economy

Extracting Distortions

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

- physical **productivity** implied by production fn.:

$$z_{f,i} = \frac{y_{f,i}}{l_{f,i}^{\gamma_i} \Pi_g \left(x_{f,g,i}^{\alpha_{g,i}}\right)}$$

- $\rightarrow$ extracted fundamentals rationalize observed dispersion b/w farms

  frictionless economy $\quad \Leftrightarrow \quad \tau_{f,g} \tau_{f,g,i} = 1$

Extracting Distortions

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

- physical **productivity** implied by production fn.:

$$z_{f,i} = \frac{y_{f,i}}{l_{f,i}^{\gamma_i} \Pi_g \left(x_{f,g,i}^{\alpha_{g,i}}\right)}$$

- $\rightarrow$ extracted fundamentals rationalize observed dispersion b/w farms

  frictionless economy $\quad \Leftrightarrow \quad \tau_{f,g} \tau_{f,g,i} = 1 \quad \Leftrightarrow \quad mrpg_{f,i} = \overline{mrpg}_i, \; mrpl_{f,i} = \overline{mrpl}_i$

Extracting Distortions

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

- physical **productivity** implied by production fn.:

$$z_{f,i} = \frac{y_{f,i}}{l_{f,i}^{\gamma_i} \Pi_g \left(x_{f,g,i}^{\alpha_{g,i}}\right)}$$

- $\rightarrow$ extracted fundamentals rationalize observed dispersion b/w farms

frictionless economy $\Leftrightarrow$ $\tau_{f,g} \tau_{f,g,i} = 1$ $\Leftrightarrow$ $mrpg_{f,i} = \overline{mrpg}_i, \; mrpl_{f,i} = \overline{mrpl}_i$

distorted economy

EXTRACTING DISTORTIONS

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

- physical **productivity** implied by production fn.:

$$z_{f,i} = \frac{y_{f,i}}{l_{f,i}^{\gamma_i} \Pi_g \left(x_{f,g,i}^{\alpha_{g,i}}\right)}$$

- $\rightarrow$ extracted fundamentals rationalize observed dispersion b/w farms

$$\text{frictionless economy} \quad \Leftrightarrow \quad \tau_{f,g}\tau_{f,g,i} = 1 \quad \Leftrightarrow \quad mrpg_{f,i} = \overline{mrpg}_i,\ mrpl_{f,i} = \overline{mrpl}_i$$

$$\text{distorted economy} \quad \Leftrightarrow \quad \text{heterog. } \tau_{f,g}\tau_{f,g,i}$$

Extracting Distortions

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

- physical **productivity** implied by production fn.:

$$z_{f,i} = \frac{y_{f,i}}{l_{f,i}^{\gamma_i} \Pi_g \left(x_{f,g,i}^{\alpha_{g,i}}\right)}$$

- $\rightarrow$ extracted fundamentals rationalize observed dispersion b/w farms

$$\text{frictionless economy} \quad \Leftrightarrow \quad \tau_{f,g}\tau_{f,g,i} = 1 \quad \Leftrightarrow \quad mrpg_{f,i} = \overline{mrpg}_i,\ mrpl_{f,i} = \overline{mrpl}_i$$

$$\text{distorted economy} \quad \Leftrightarrow \quad \text{heterog. } \tau_{f,g}\tau_{f,g,i} \quad \Leftrightarrow \quad \text{heterog. } mrpg_{f,i},\ mrpl_{f,i}$$

Extracting Distortions

- unobserved **distortions** map to observed **marginal revenue products**:

$$r_g \tau_{f,g} \tau_{f,g,i} = \frac{\alpha_{g,i} \eta \left(p_i y_{f,i}\right)^{\eta}}{x_{f,g,i}} = mrpg_{f,i}$$

$$\lambda_f \tau_{f,l,i} = \frac{\gamma_i \eta \left(p_i y_{f,i}\right)^{\eta}}{l_{f,i}} = mrpl_{f,i}$$

- physical **productivity** implied by production fn.:

$$z_{f,i} = \frac{y_{f,i}}{l_{f,i}^{\gamma_i} \Pi_g \left(x_{f,g,i}^{\alpha_{g,i}}\right)}$$

- $\rightarrow$ extracted fundamentals rationalize observed dispersion b/w farms

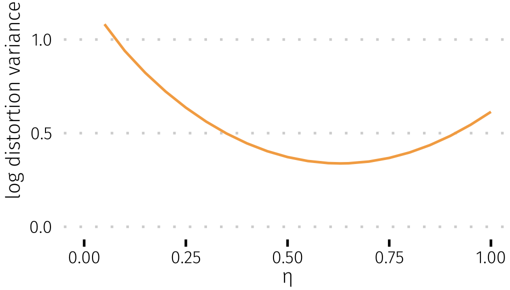
$$\text{frictionless economy} \quad \Leftrightarrow \quad \tau_{f,g}\tau_{f,g,i} = 1 \quad \Leftrightarrow \quad mrpg_{f,i} = \overline{mrpg}_i, \; mrpl_{f,i} = \overline{mrpl}_i$$

$$\text{distorted economy} \quad \Leftrightarrow \quad \text{heterog. } \tau_{f,g}\tau_{f,g,i} \quad \Leftrightarrow \quad \text{heterog. } mrpg_{f,i}, \; mrpl_{f,i}$$

▶ *splitting distortions*

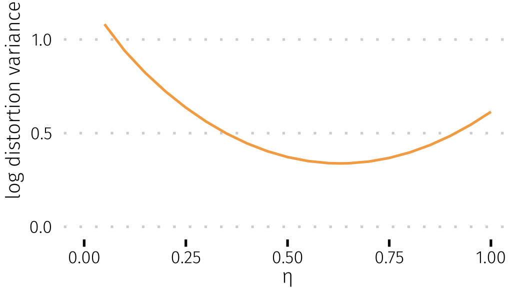
Choose η That Minimizes Implied Dispersion

- τs reproducing data need to be extreme if η is too high or too low



Choose η That Minimizes Implied Dispersion

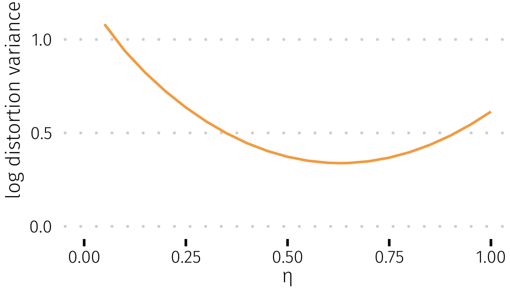
- τs reproducing data need to be extreme if η is too high or too low



- **low** η: farm, farm-crop output more uniformly distributed in efficient allocation
  $\rightarrow$ data farm size **“too varied”**, farms mix crops **“too little”** $\rightarrow$ extreme distortions

Choose η That Minimizes Implied Dispersion

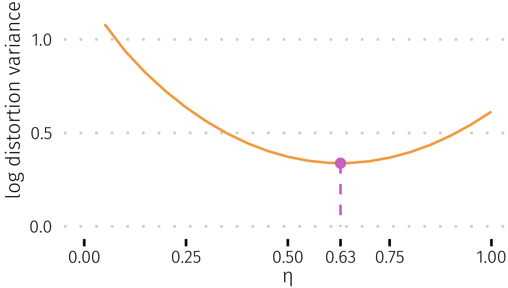
- τs reproducing data need to be extreme if η is too high or too low



- **low** η: farm, farm-crop output more uniformly distributed in efficient allocation
  - $\rightarrow$ data farm size **"too varied"**, farms mix crops **"too little"** $\rightarrow$ extreme distortions
- **high** η: farm, farm-crop output more dispersed in efficient allocation
  - $\rightarrow$ data farm size **"too uniform"**, farms mix crops **"too much"** $\rightarrow$ extreme distortions

Choose η That Minimizes Implied Dispersion

- τs reproducing data need to be extreme if η is too high or too low



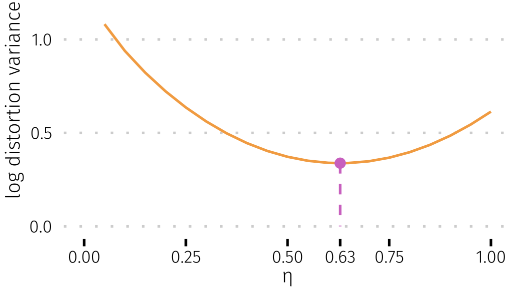
- **low** η: farm, farm-crop output more uniformly distributed in efficient allocation
  - $\rightarrow$ data farm size **"too varied"**, farms mix crops **"too little"** $\rightarrow$ extreme distortions
- **high** η: farm, farm-crop output more dispersed in efficient allocation
  - $\rightarrow$ data farm size **"too uniform"**, farms mix crops **"too much"** $\rightarrow$ extreme distortions

$\rightarrow$ pick η that minimizes distortions required to explain observed output distribution

$\rightarrow$ conservative misallocation estimates

Choose η That Minimizes Implied Dispersion

- τs reproducing data need to be extreme if η is too high or too low



- **low** η: farm, farm-crop output more uniformly distributed in efficient allocation
  - $\rightarrow$ data farm size **"too varied"**, farms mix crops **"too little"** $\rightarrow$ extreme distortions
- **high** η: farm, farm-crop output more dispersed in efficient allocation
  - $\rightarrow$ data farm size **"too uniform"**, farms mix crops **"too much"** $\rightarrow$ extreme distortions

$\rightarrow$ pick η that minimizes distortions required to explain observed output distribution

$\rightarrow$ conservative misallocation estimates

▶ *details*

Reallocation

- to quantify **aggregate cost** of distortions, conduct **counterfactual reallocations**

Reallocation

- to quantify **aggregate cost** of distortions, conduct **counterfactual reallocations**
- equalize (or reduce) distortions between farms

Reallocation

- to quantify **aggregate cost** of distortions, conduct **counterfactual reallocations**
- equalize (or reduce) distortions between farms
- prohibit product set switching

Reallocation

- to quantify **aggregate cost** of distortions, conduct **counterfactual reallocations**
- equalize (or reduce) distortions between farms
- prohibit product set switching
  - $\rightarrow$ don't need productivities, frictions for unobserved farm-product combinations

Reallocation

- to quantify **aggregate cost** of distortions, conduct **counterfactual reallocations**
- equalize (or reduce) distortions between farms
- prohibit product set switching
  - $\rightarrow$ don't need productivities, frictions for unobserved farm-product combinations
- treat each season as a separate economy
  - to prevent double-booking of inputs in a single season
  - sum aggregate output across seasons

Reallocation

- to quantify **aggregate cost** of distortions, conduct **counterfactual reallocations**
- equalize (or reduce) distortions between farms
- prohibit product set switching
  - $\rightarrow$ don't need productivities, frictions for unobserved farm-product combinations
- treat each season as a separate economy
  - to prevent double-booking of inputs in a single season
  - sum aggregate output across seasons
- compute counterfactual Δoutput

▶ *reallocation exercise details*

Reallocation

- to quantify **aggregate cost** of distortions, conduct **counterfactual reallocations**
- equalize (or reduce) distortions between farms
- prohibit product set switching
  - → don't need productivities, frictions for unobserved farm-product combinations
- treat each season as a separate economy
  - to prevent double-booking of inputs in a single season
  - sum aggregate output across seasons
- compute counterfactual Δoutput
  - **reallocation gain** = Δoutput between counterfactual and current

▶ *reallocation exercise details*

Reallocation

- to quantify **aggregate cost** of distortions, conduct **counterfactual reallocations**
- equalize (or reduce) distortions between farms
- prohibit product set switching
  - $\rightarrow$ don't need productivities, frictions for unobserved farm-product combinations
- treat each season as a separate economy
  - to prevent double-booking of inputs in a single season
  - sum aggregate output across seasons
- compute counterfactual Δoutput
  - **reallocation gain** = Δoutput between counterfactual and current
  - aggregate inputs are fixed $\rightarrow$ aggregate output $\uparrow$ reflects aggregate TFP $\uparrow$

▶ *reallocation exercise details*

Reallocation

- to quantify **aggregate cost** of distortions, conduct **counterfactual reallocations**
- equalize (or reduce) distortions between farms
- prohibit product set switching
  $\rightarrow$ don't need productivities, frictions for unobserved farm-product combinations
- treat each season as a separate economy
  - to prevent double-booking of inputs in a single season
  - sum aggregate output across seasons
- compute counterfactual Δoutput
  - **reallocation gain** = Δoutput between counterfactual and current
  - aggregate inputs are fixed $\rightarrow$ aggregate output $\uparrow$ reflects aggregate TFP $\uparrow$
  - farm-level TFPs are fixed $\rightarrow$ aggregate TFP $\uparrow$ reflects pure **reallocation gain** or **misallocation cost**

▶ *reallocation exercise details*

Exercises

Benchmark Exercise

reallocation gain

Exercises

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$

reallocation gain

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$
- keep within-farm distortions $\tau_{f,l,i}, \tau_{f,g,i}$

reallocation gain

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$
- keep within-farm distortions $\tau_{f,l,i}, \tau_{f,g,i}$
  - preserve idiosyncratic product choice motives

reallocation gain

Exercises

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$
- keep within-farm distortions $\tau_{f,l,i}, \tau_{f,g,i}$
  - preserve idiosyncratic product choice motives

reallocation gain

multi-product:	294%	
1-product:	212%	(28% $\downarrow$)

Exercises

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$
- keep within-farm distortions $\tau_{f,l,i}, \tau_{f,g,i}$
  - preserve idiosyncratic product choice motives

Match Least-Distorted State

reallocation gain

multi-product:	294%	
1-product:	212%	(28% $\downarrow$)

Exercises

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$
- keep within-farm distortions $\tau_{f,l,i}, \tau_{f,g,i}$
  - preserve idiosyncratic product choice motives

Match Least-Distorted State

- more conservative and practical

	reallocation gain	
multi-product:	294%	
1-product:	212%	(28% ↓)

Exercises

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$
- keep within-farm distortions $\tau_{f,l,i}, \tau_{f,g,i}$
  - preserve idiosyncratic product choice motives

Match Least-Distorted State

- more conservative and practical
- downscale frictions s.t. their variances match those in the least-distorted state

	reallocation gain	
multi-product:	294%	
1-product:	212%	(28% ↓)

Exercises

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$
- keep within-farm distortions $\tau_{f,l,i}, \tau_{f,g,i}$
  - preserve idiosyncratic product choice motives

Match Least-Distorted State

- more conservative and practical
- downscale frictions s.t. their variances match those in the least-distorted state
  - Tamil Nadu has the least dispersed distortions

reallocation gain

multi-product:	294%	
1-product:	212%	(28% ↓)

Exercises

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$
- keep within-farm distortions $\tau_{f,l,i}, \tau_{f,g,i}$
  - preserve idiosyncratic product choice motives

Match Least-Distorted State

- more conservative and practical
- downscale frictions s.t. their variances match those in the least-distorted state
  - Tamil Nadu has the least dispersed distortions

reallocation gain

multi-product:	294%	
1-product:	212%	(28% ↓)

multi-product:	35%	
1-product:	45%	(26% ↑)

Exercises

Benchmark Exercise

- equalize cross-farm distortions $\lambda_f, \tau_{f,g}$
- keep within-farm distortions $\tau_{f,l,i}, \tau_{f,g,i}$
  - preserve idiosyncratic product choice motives

reallocation gain

multi-product: 294%
1-product: 212% (28% ↓)

Match Least-Distorted State

- more conservative and practical
- downscale frictions s.t. their variances match those in the least-distorted state
  - Tamil Nadu has the least dispersed distortions

multi-product: 35%
1-product: 45% (26% ↑)

▶ *sensitivity to concavity* ▶ *role of states, seasons*

Mechanisms Contributing to 1-Product Model Error

Mechanism I: Product Heterogeneity as Frictions

- 1-product model misinterprets crop heterogeneity as frictions

Mechanism I: Product Heterogeneity as Frictions

- 1-product model misinterprets crop heterogeneity as frictions
  - farm 1 draws high z in rice farm 2 draws high z in vegetables

Mechanism I: Product Heterogeneity as Frictions

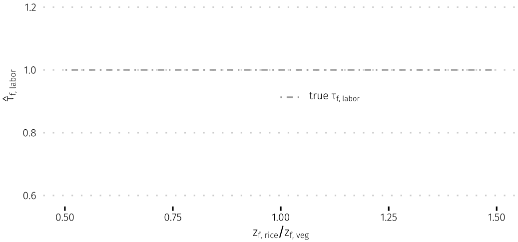
- 1-product model misinterprets crop heterogeneity as frictions
  - farm 1 draws high z in rice farm 2 draws high z in vegetables
    - farm 1: high $\frac{y_{1,rice}}{y_{1,veg}}$ $\rightarrow$ high $\frac{l_1}{x_{1,labor}}$

Mechanism I: Product Heterogeneity as Frictions

- 1-product model misinterprets crop heterogeneity as frictions
  - farm 1 draws high z in rice farm 2 draws high z in vegetables
    - farm 1: high $\frac{y_{1,rice}}{y_{1,veg}} \rightarrow$ high $\frac{l_1}{x_{1,labor}}$
    - farm 2: low $\frac{y_{2,rice}}{y_{2,veg}} \rightarrow$ low $\frac{l_2}{x_{2,labor}}$

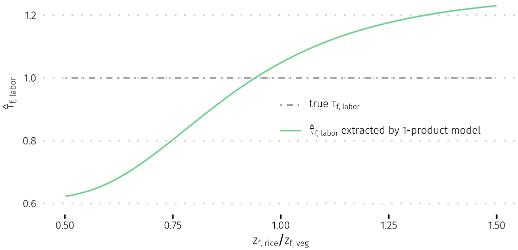
Mechanism I: Product Heterogeneity as Frictions

- 1-product model misinterprets crop heterogeneity as frictions
  - farm 1 draws high z in rice farm 2 draws high z in vegetables
    - farm 1: high $\frac{y_{1,rice}}{y_{1,veg}} \rightarrow$ high $\frac{l_1}{x_{1,labor}}$
    - farm 2: low $\frac{y_{2,rice}}{y_{2,veg}} \rightarrow$ low $\frac{l_2}{x_{2,labor}}$
  - assume no frictions $\rightarrow \frac{\alpha_{g,rice}\eta(p_{rice}y_1)^\eta}{x_{1,g}} = \frac{\alpha_{labor,veg}\eta(p_{veg}y_2)^\eta}{x_{2,labor}}$ in multi-crop model



Mechanism I: Product Heterogeneity as Frictions

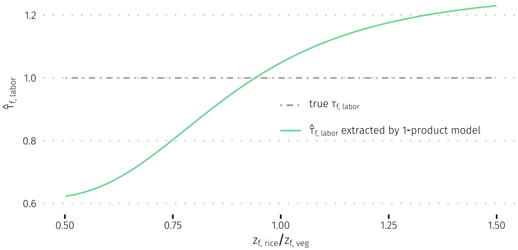
- 1-product model misinterprets crop heterogeneity as frictions
  - farm 1 draws high z in rice farm 2 draws high z in vegetables
    - farm 1: high $\frac{y_{1,rice}}{y_{1,veg}} \rightarrow$ high $\frac{l_1}{x_{1,labor}}$
    - farm 2: low $\frac{y_{2,rice}}{y_{2,veg}} \rightarrow$ low $\frac{l_2}{x_{2,labor}}$
  - assume no frictions $\rightarrow \frac{\alpha_{g,rice}\eta(p_{rice}y_1)^\eta}{x_{1,g}} = \frac{\alpha_{labor,veg}\eta(p_{veg}y_2)^\eta}{x_{2,labor}}$ in multi-crop model
  - 1-product model: $\frac{\alpha_{labor,agg}\eta(p_{agg}y_1)^\eta}{x_{1,labor}} \neq \frac{\alpha_{labor,agg}\eta(p_{agg}y_2)^\eta}{x_{2,labor}} \rightarrow$ imputes frictions



Mechanism I: Product Heterogeneity as Frictions

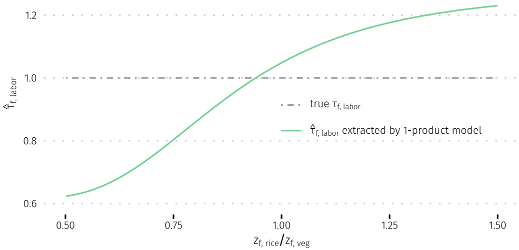
- 1-product model misinterprets crop heterogeneity as frictions
  - farm 1 draws high z in rice farm 2 draws high z in vegetables
    - farm 1: high $\frac{y_{1,rice}}{y_{1,veg}} \rightarrow$ high $\frac{l_1}{x_{1,labor}}$
    - farm 2: low $\frac{y_{2,rice}}{y_{2,veg}} \rightarrow$ low $\frac{l_2}{x_{2,labor}}$
  - assume no frictions $\rightarrow \frac{\alpha_{g,rice}\eta(p_{rice}y_1)^\eta}{x_{1,g}} = \frac{\alpha_{labor,veg}\eta(p_{veg}y_2)^\eta}{x_{2,labor}}$ in multi-crop model
  - 1-product model: $\frac{\alpha_{labor,agg}\eta(p_{agg}y_1)^\eta}{x_{1,labor}} \neq \frac{\alpha_{labor,agg}\eta(p_{agg}y_2)^\eta}{x_{2,labor}} \rightarrow$ imputes frictions

$\rightarrow$ 1-product model overstates **misallocation**



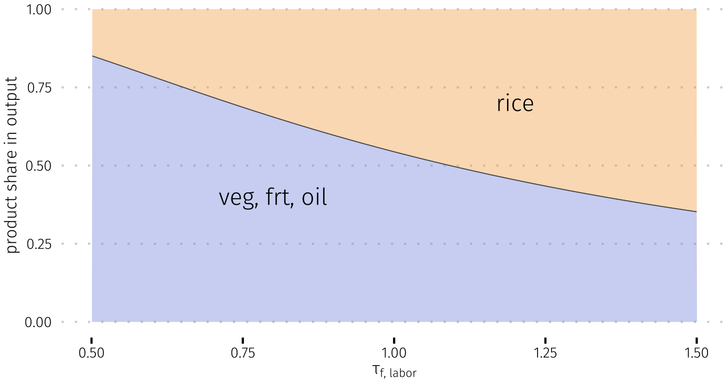
Mechanism I: Product Heterogeneity as Frictions

- 1-product model misinterprets crop heterogeneity as frictions
  - farm 1 draws high z in rice farm 2 draws high z in vegetables
    - farm 1: high $\frac{y_{1,rice}}{y_{1,veg}} \rightarrow$ high $\frac{l_1}{x_{1,labor}}$
    - farm 2: low $\frac{y_{2,rice}}{y_{2,veg}} \rightarrow$ low $\frac{l_2}{x_{2,labor}}$
  - assume no frictions $\rightarrow \frac{\alpha_{g,rice}\eta(p_{rice}y_1)^\eta}{x_{1,g}} = \frac{\alpha_{labor,veg}\eta(p_{veg}y_2)^\eta}{x_{2,labor}}$ in multi-crop model
  - 1-product model: $\frac{\alpha_{labor,agg}\eta(p_{agg}y_1)^\eta}{x_{1,labor}} \neq \frac{\alpha_{labor,agg}\eta(p_{agg}y_2)^\eta}{x_{2,labor}} \rightarrow$ imputes frictions
  - $\rightarrow$ 1-product model overstates **misallocation**
- **exercise to isolate**: apply 1-product model to counterfactual reallocation data generated by multi-product model



Mechanism II: Endogenous Product Choice

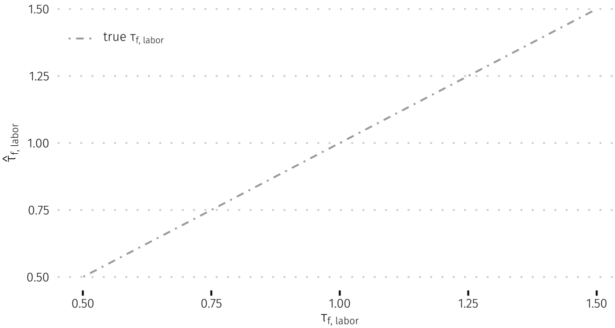
- simulated multi-product farm as labor distortion $\tau_{f,labor}$ is varied:



- $\tau_{f,labor} \uparrow \quad \rightarrow$ shift from labor-intensive vegetables to land-intensive rice

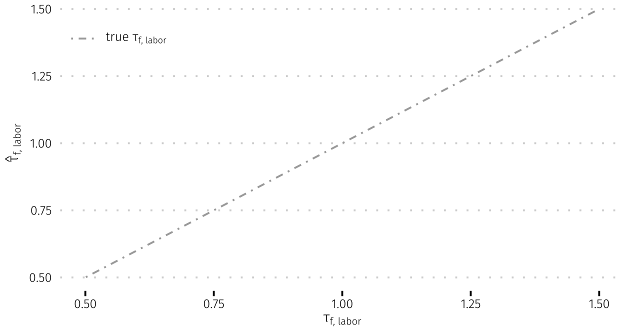
Mechanism II: Endogenous Product Choice

- apply single-product model to extract frictions from simulated multi-product data



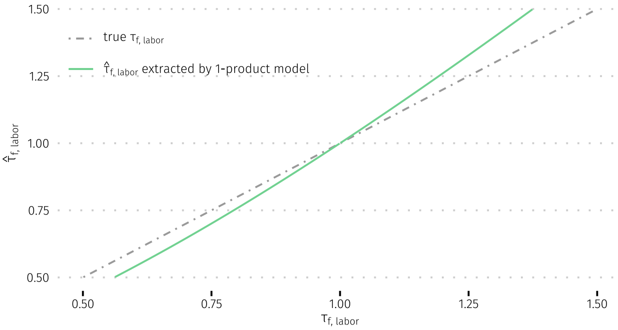
Mechanism II: Endogenous Product Choice

- apply single-product model to extract frictions from simulated multi-product data
- multi-product model: optimal product choice response to frictions
  - modest $\tau_{f,labor}$ increase $\rightarrow$ shift to land-intensive rice $\rightarrow$ hire even more land rel. to labor



Mechanism II: Endogenous Product Choice

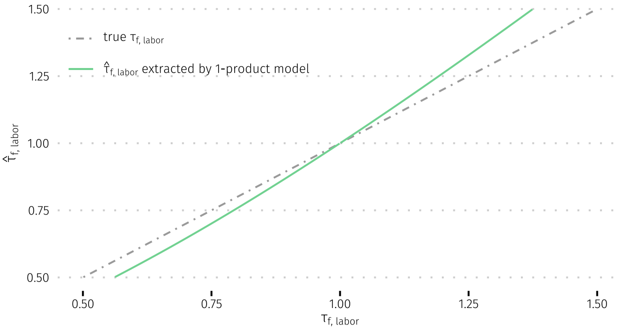
- apply single-product model to extract frictions from simulated multi-product data
- multi-product model: optimal product choice response to frictions
  - modest $\tau_{f,labor}$ increase $\rightarrow$ shift to land-intensive rice $\rightarrow$ hire even more land rel. to labor
- 1-product model: high input ratio dispersion $\rightarrow$ infer large heterogeneity in frictions



Mechanism II: Endogenous Product Choice

- apply single-product model to extract frictions from simulated multi-product data
- multi-product model: optimal product choice response to frictions
  - modest $\tau_{f,labor}$ increase $\rightarrow$ shift to land-intensive rice $\rightarrow$ hire even more land rel. to labor
- 1-product model: high input ratio dispersion $\rightarrow$ infer large heterogeneity in frictions

$\rightarrow$ 1-crop model overstates **misallocation**



Mechanism II: Endogenous Product Choice

- **exercise to isolate**: prohibit farms in multi-product model to change product choice in counterfactuals
  - keep product sets fixed
  - keep input allocation across crops fixed: farm can choose $\sum_i x_{f,g,i}$ but $x_{f,g,i}$ gets a fixed share of total

Mechanism III: TFP Dispersion

- 1-product model understates TFP dispersion

Mechanism III: TFP Dispersion

- 1-product model understates TFP dispersion
  - a farm has z_L in rice, z_H in vegetables

Mechanism III: TFP Dispersion

- 1-product model understates TFP dispersion
  - a farm has z_L in rice, z_H in vegetables
  - severe frictions $\rightarrow$ farm mostly grows rice $\rightarrow$ avg TFP $= z_M < z_H$

Mechanism III: TFP Dispersion

- 1-product model understates TFP dispersion
  - a farm has z_L in rice, z_H in vegetables
  - severe frictions $\rightarrow$ farm mostly grows rice $\rightarrow$ avg TFP $= z_M < z_H$
  - multi-crop model: remove frictions $\rightarrow$ farm switches to vegetables $\rightarrow$ avg TFP $= z_H$

Mechanism III: TFP Dispersion

- 1-product model understates TFP dispersion
  - a farm has z_L in rice, z_H in vegetables
  - severe frictions $\rightarrow$ farm mostly grows rice $\rightarrow$ avg TFP $= z_M < z_H$
  - multi-crop model: remove frictions $\rightarrow$ farm switches to vegetables $\rightarrow$ avg TFP $= z_H$
  - 1-product model: remove frictions $\rightarrow$ avg TFP still z_M

Mechanism III: TFP Dispersion

- 1-product model understates TFP dispersion
  - a farm has z_L in rice, z_H in vegetables
  - severe frictions $\rightarrow$ farm mostly grows rice $\rightarrow$ avg TFP $= z_M < z_H$
  - multi-crop model: remove frictions $\rightarrow$ farm switches to vegetables $\rightarrow$ avg TFP $= z_H$
  - 1-product model: remove frictions $\rightarrow$ avg TFP still z_M

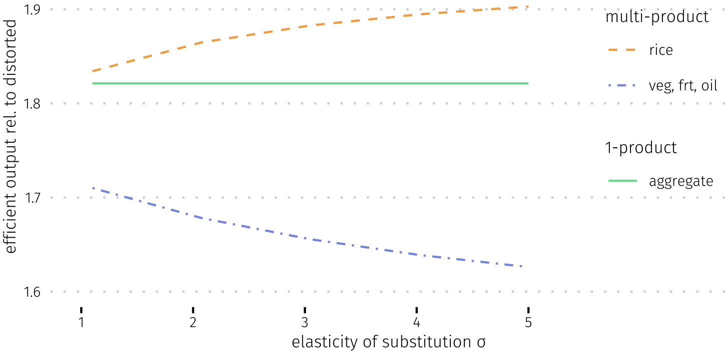
  $\rightarrow$ 1-product model understates **misallocation**

Mechanism III: TFP Dispersion

- 1-product model understates TFP dispersion
  - a farm has z_L in rice, z_H in vegetables
  - severe frictions $\rightarrow$ farm mostly grows rice $\rightarrow$ avg TFP $= z_M < z_H$
  - multi-crop model: remove frictions $\rightarrow$ farm switches to vegetables $\rightarrow$ avg TFP $= z_H$
  - 1-product model: remove frictions $\rightarrow$ avg TFP still z_M
  - $\rightarrow$ 1-product model understates **misallocation**
- **exercise to isolate**: treat farm-crops as separate farms for 1-product model

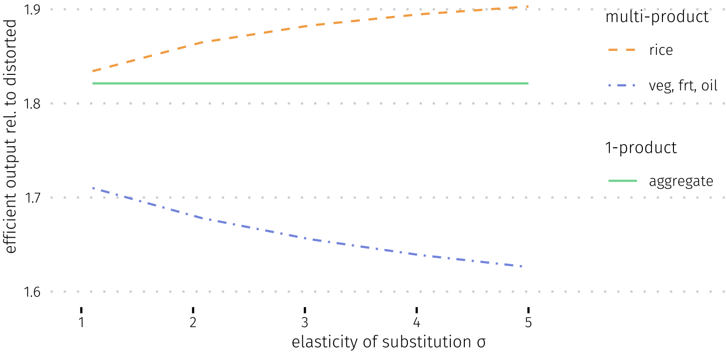
Mechanism IV: Returns to Scale

- some products have higher returns to scale → some farms grow more in reallocation



Mechanism IV: Returns to Scale

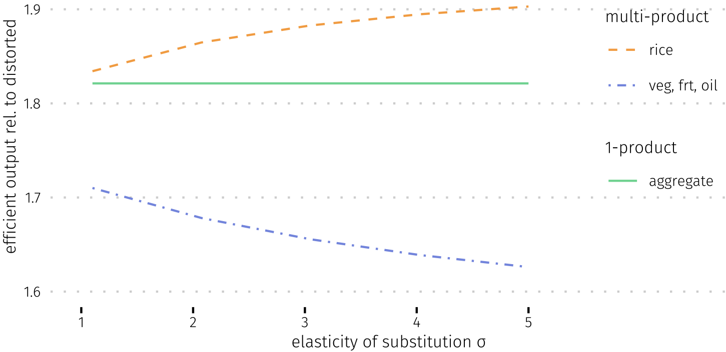
- some products have higher returns to scale → some farms grow more in reallocation
- → consumer can substitute toward high-RS products to take advantage



Mechanism IV: Returns to Scale

- some products have higher returns to scale → some farms grow more in reallocation
- → consumer can substitute toward high-RS products to take advantage

→ 1-product model understates **misallocation**

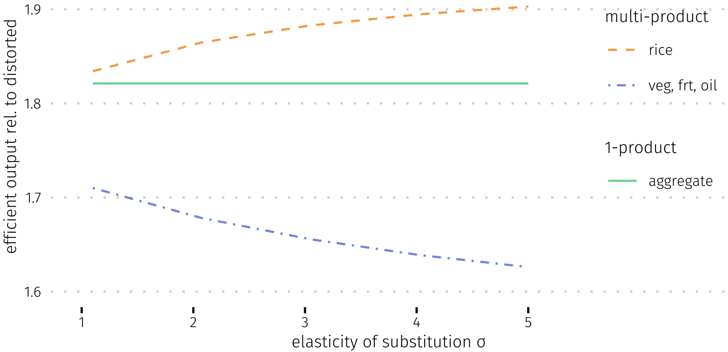


Mechanism IV: Returns to Scale

- some products have higher returns to scale → some farms grow more in reallocation
- → consumer can substitute toward high-RS products to take advantage

→ 1-product model understates **misallocation**

- **exercise to isolate**: rescale input elasticities to equalize returns to scale

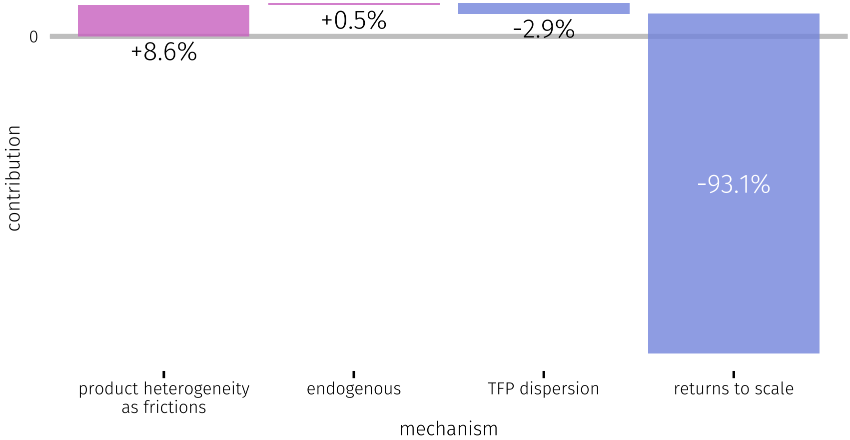


Mechanisms Decomposition: Benchmark Reallocation

- **benchmark**: single-product model understates gain by **82 pp** (28%)

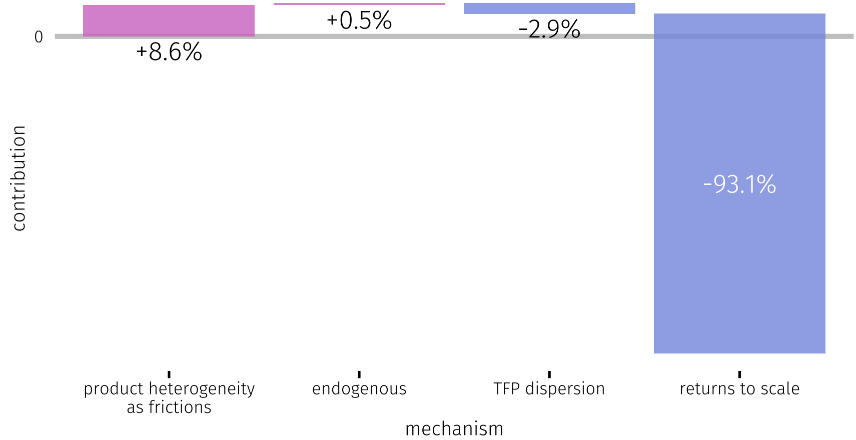
Mechanisms Decomposition: Benchmark Reallocation

- **benchmark**: single-product model understates gain by **82 pp** (28%)



Mechanisms Decomposition: Benchmark Reallocation

- **benchmark**: single-product model understates gain by **82 pp** (28%)



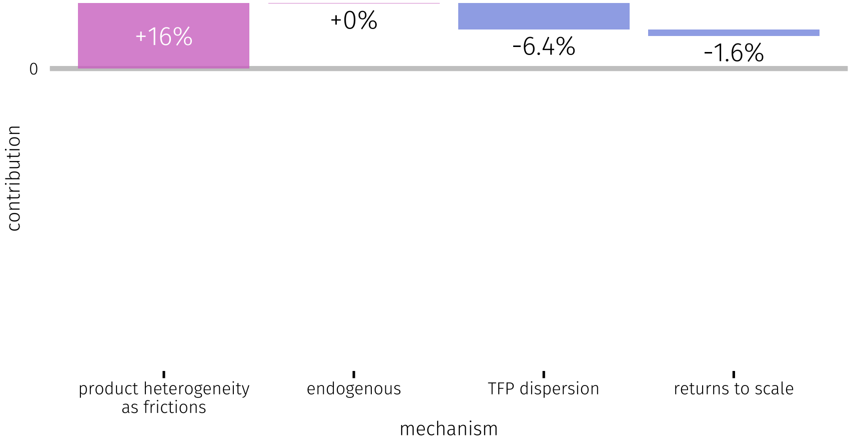
- assess total drag of misallocation → farms' ability to expand matters most
  → single-product model understates **misallocation**

Mechanisms Decomposition: "Least-Distorted State" Reallocation

- **"least-distorted state"**: single-product model overstates gain by **10 pp** (26%)

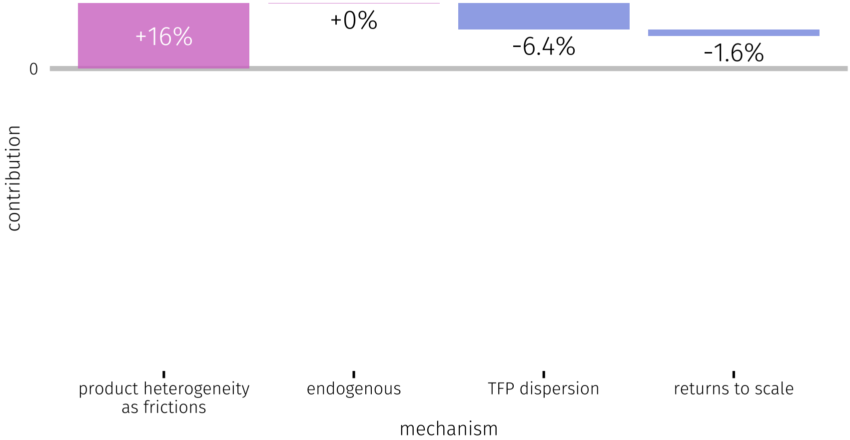
Mechanisms Decomposition: “Least-Distorted State” Reallocation

- **“least-distorted state”**: single-product model overstates gain by **10 pp** (26%)



Mechanisms Decomposition: "Least-Distorted State" Reallocation

- **"least-distorted state"**: single-product model overstates gain by **10 pp** (26%)



- consider partial reallocations → estimation of frictions matters most
  → single-product model overstates **misallocation** ▶ *details*

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India
- multi-product model shows that conventional 1-product models:

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India
- multi-product model shows that conventional 1-product models:
  - overstate misallocation by

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India
- multi-product model shows that conventional 1-product models:
  - overstate misallocation by
    - > misinterpreting product heterogeneity as frictions

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India
- multi-product model shows that conventional 1-product models:
  - overstate misallocation by
    - > misinterpreting product heterogeneity as frictions
    - > missing endogenous product choice response

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India
- multi-product model shows that conventional 1-product models:
  - overstate misallocation by
    - > misinterpreting product heterogeneity as frictions
    - > missing endogenous product choice response
  - understate misallocation by

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India
- multi-product model shows that conventional 1-product models:
  - overstate misallocation by
    - misinterpreting product heterogeneity as frictions
    - missing endogenous product choice response
  - understate misallocation by
    - ignoring within-farm productivity heterogeneity

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India
- multi-product model shows that conventional 1-product models:
  - overstate misallocation by
    - misinterpreting product heterogeneity as frictions
    - missing endogenous product choice response
  - understate misallocation by
    - ignoring within-farm productivity heterogeneity
    - ignoring returns-to-scale heterogeneity across crops

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India
- multi-product model shows that conventional 1-product models:
  - overstate misallocation by
    - misinterpreting product heterogeneity as frictions
    - missing endogenous product choice response
  - understate misallocation by
    - ignoring within-farm productivity heterogeneity
    - ignoring returns-to-scale heterogeneity across crops
- benchmark exercise: 1-product model understates misallocation

Conclusion

- **existing misallocation estimates**: single product or multiple homogeneous products
- estimate heterogeneous production functions across crops in India
- multi-product model shows that conventional 1-product models:
  - overstate misallocation by
    - > misinterpreting product heterogeneity as frictions
    - > missing endogenous product choice response
  - understate misallocation by
    - > ignoring within-farm productivity heterogeneity
    - > ignoring returns-to-scale heterogeneity across crops
- benchmark exercise: 1-product model understates misallocation
  - but overstates what can be attained with partial policies

Next

- *WIP*: input-driven production method heterogeneity and misallocation

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features
  - heterogeneity in land features $\rightarrow$ different production methods

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features
  - heterogeneity in land features → different production methods
  - production method heterogeneity can get mislabeled as misallocation

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features
  - heterogeneity in land features $\rightarrow$ different production methods
  - production method heterogeneity can get mislabeled as misallocation
- *WIP*: gains from trade when domestic trade is costly, input markets distorted

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features
  - heterogeneity in land features $\rightarrow$ different production methods
  - production method heterogeneity can get mislabeled as misallocation
- *WIP*: gains from trade when domestic trade is costly, input markets distorted
  - **empirics**: farmers' response to trade liberalization depending on market access, misallocation

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features
  - heterogeneity in land features $\rightarrow$ different production methods
  - production method heterogeneity can get mislabeled as misallocation
- *WIP*: gains from trade when domestic trade is costly, input markets distorted
  - **empirics**: farmers' response to trade liberalization depending on market access, misallocation
  - **model**: spatial Ricardian model + misallocation

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features
  - heterogeneity in land features $\rightarrow$ different production methods
  - production method heterogeneity can get mislabeled as misallocation
- *WIP*: gains from trade when domestic trade is costly, input markets distorted
  - **empirics**: farmers' response to trade liberalization depending on market access, misallocation
  - **model**: spatial Ricardian model + misallocation
- study the dynamics of product choice

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features
  - heterogeneity in land features $\rightarrow$ different production methods
  - production method heterogeneity can get mislabeled as misallocation
- *WIP*: gains from trade when domestic trade is costly, input markets distorted
  - **empirics**: farmers' response to trade liberalization depending on market access, misallocation
  - **model**: spatial Ricardian model + misallocation
- study the dynamics of product choice
  - crop rotation, crop complimentarities between seasons, years

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features
  - heterogeneity in land features $\rightarrow$ different production methods
  - production method heterogeneity can get mislabeled as misallocation
- *WIP*: gains from trade when domestic trade is costly, input markets distorted
  - **empirics**: farmers' response to trade liberalization depending on market access, misallocation
  - **model**: spatial Ricardian model + misallocation
- study the dynamics of product choice
  - crop rotation, crop complimentarities between seasons, years
  - sticky input choices between seasons

Next

- *WIP*: input-driven production method heterogeneity and misallocation
  - cluster plots into land types by physical land features
  - heterogeneity in land features $\rightarrow$ different production methods
  - production method heterogeneity can get mislabeled as misallocation
- *WIP*: gains from trade when domestic trade is costly, input markets distorted
  - **empirics**: farmers' response to trade liberalization depending on market access, misallocation
  - **model**: spatial Ricardian model + misallocation
- study the dynamics of product choice
  - crop rotation, crop complimentarities between seasons, years
  - sticky input choices between seasons

  $\rightarrow$ input choices may appear inefficient statically but be optimal dynamically

Farm Solution Expressions

$$\sum_{i\in I_f}\left(\lambda_f^{\frac{\eta\sum_g \alpha_{g,i}-1}{1-\eta\sum_g \alpha_{g,i}-\eta\gamma_i}}\right)\left(\left(p_i z_{f,i}\right)^{\eta}\eta\left(\frac{\gamma_i}{\tau_{f,l,i}}\right)^{1-\eta\sum_g \alpha_{g,i}}\Pi_g\left(\frac{\alpha_{g,i}}{r_g\tau_{f,g}\tau_{f,g,i}}\right)^{\eta\alpha_{g,i}}\right)^{\frac{1}{1-\eta\sum_g \alpha_{g,i}-\eta\gamma_i}}\tau_{f,l,i}=L_f$$

$$x_{f,g,i}=\frac{\alpha_{g,i}}{r_g\tau_{f,g}\tau_{f,g,i}}\left(\frac{\gamma_i}{\lambda_f\tau_{f,l,i}}\right)^{\frac{\eta\gamma_i}{1-\eta\sum_g \alpha_{g,i}-\eta\gamma_i}}\left(\left(p_i z_{f,i}\right)^{\eta}\eta\Pi_h\left(\frac{\alpha_{h,i}}{r_h\tau_{f,h}\tau_{f,h,i}}\right)^{\eta\alpha_{h,i}}\right)^{\frac{1}{1-\eta\sum_h \alpha_{h,i}-\eta\gamma_i}}$$

$$l_{f,i}=\left(\left(p_i z_{f,i}\right)^{\eta}\eta\left(\frac{\gamma_i}{\lambda_f\tau_{f,l,i}}\right)^{1-\eta\sum_g \alpha_{g,i}}\Pi_g\left(\frac{\alpha_{g,i}}{r_g\tau_{f,g}\tau_{f,g,i}}\right)^{\eta\alpha_{g,i}}\right)^{\frac{1}{1-\eta\sum_g \alpha_{g,i}-\eta\gamma_i}}$$

back

List of crops

Crop list

Rice	Wheat	Other Cereals	Pulses	Oilseeds, Fruits and Vegetables		
				Oilseeds	**Vegetables**	**Fruits /Condiments**
		Barley	Black gram	Sesame	Ash gourd	Mango
		Maize	Green peas	Groundnut	Beet root	Papaya
		Sorghum	Pigeon peas	Castor	Bitter gourd	Grapes
		Pearl millet	Horse gram	Sunflower	Bottle gourd	Plum
		Finger millet	Cowpea	Niger	Eggplant	Cardamom
		Others	Kidney bean	Soybean	Board bean	Chilli
			Lentil	Safflower	Cabbage	Cumin
			Chickpeas	Rapseed	Cauliflower	Dill seed
			Others	Linseed	Carrot	Indian mustard
				Others	Potato	Other
					Cucumber	
					Peas	

back

Inputs, Outputs

- **labor**:
  - family and hired labor days by plots

back

Inputs, Outputs

- **labor**:
  - family and hired labor days by plots
- **land**:
  - quality-adjusted land measure

back

Inputs, Outputs

- **labor**:
  - family and hired labor days by plots
- **land**:
  - quality-adjusted land measure
- **intermediate inputs**:
  - expenses on seeds, fertilizer, irrigation, machinery and bullocks, and fuels

back

Inputs, Outputs

- **labor**:
  - family and hired labor days by plots
- **land**:
  - quality-adjusted land measure
- **intermediate inputs**:
  - expenses on seeds, fertilizer, irrigation, machinery and bullocks, and fuels
- **output**:
  - market value of quantity harvested

back

Instrumental Variables

- assumptions:

back

Instrumental Variables

- **assumptions**:
  - **observed** shocks to farm f's plot k affect **shadow price** of inputs on f's plot j

back

Instrumental Variables

- **assumptions**:
  - **observed** shocks to farm f's plot k affect **shadow price** of inputs on f's plot j
  - **observed** shocks affecting input demand on plot k are not correlated with **unobserved** shocks affecting input demand on plot j, conditional on observed shocks to j

back

Instrumental Variables

- **assumptions**:
  - **observed** shocks to farm f's plot k affect **shadow price** of inputs on f's plot j
  - **observed** shocks affecting input demand on plot k are not correlated with **unobserved** shocks affecting input demand on plot j, conditional on observed shocks to j
  - f-level shocks interacted with plot-level features provide such shocks

back

Instrumental Variables

- **assumptions**:
  - **observed** shocks to farm f's plot k affect **shadow price** of inputs on f's plot j
  - **observed** shocks affecting input demand on plot k are not correlated with **unobserved** shocks affecting input demand on plot j, conditional on observed shocks to j
  - f-level shocks interacted with plot-level features provide such shocks
- **2SLS first stage**:

$$M_{j,i,t} = Z_{k \neq j,i,t} + \mu_{j,i,t}$$

  - M = land, labor, intermediates
  - $Z_{k \neq j,i,t}$: instruments **from other plots within farm**
    - > agricultural shocks interacted with plot characteristics
    - > household, community characteristics & shocks interacted with plot characteristics

back

Inputs: Quality-Adjusted Land

- procedure

back

Inputs: Quality-Adjusted Land

- procedure
    1. split data into **training** sample (2/3) and **test** sample (1/3)

back

Inputs: Quality-Adjusted Land

- procedure
  1. split data into **training** sample (2/3) and **test** sample (1/3)
  2. estimate OLS & RF on **training** data: **reg log price/acre on land features**
     - OLS includes all 2-way interactions
     - RF tuned with k-fold cross-validation

back

Inputs: Quality-Adjusted Land

- procedure
  1. split data into **training** sample (2/3) and **test** sample (1/3)
  2. estimate OLS & RF on **training** data: **reg log price/acre on land features**
     - > OLS includes all 2-way interactions
     - > RF tuned with k-fold cross-validation
  3. compute Mean Squared Error and R^2 on **test** data
     - > to preclude overfitting

back

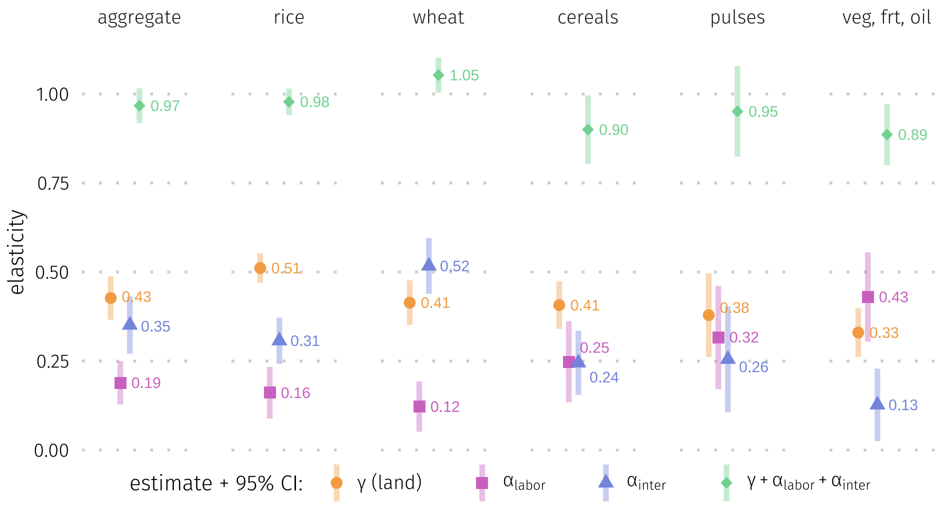
Inputs: Quality-Adjusted Land

- procedure
  1. split data into **training** sample (2/3) and **test** sample (1/3)
  2. estimate OLS & RF on **training** data: **reg log price/acre on land features**
     - > OLS includes all 2-way interactions
     - > RF tuned with k-fold cross-validation
  3. compute Mean Squared Error and R^2 on **test** data
     - > to preclude overfitting
- results

	OLS	RF
MSE	0.61	0.49
R^2	0.39	0.51

back

Production Functions: All Crops



back

Production Functions: Table

	Aggregate	Rice	Wheat	Other Cereals	Pulses	Veg, Frt, Oil
Land	0.427	0.511	0.414	0.407	0.379	0.330
	(0.031)	(0.021)	(0.032)	(0.034)	(0.060)	(0.035)
Labor	0.189	0.161	0.122	0.248	0.316	0.430
	(0.031)	(0.037)	(0.036)	(0.058)	(0.074)	(0.064)
Intermediates	0.351	0.307	0.517	0.245	0.255	0.127
	(0.041)	(0.033)	(0.040)	(0.046)	(0.076)	(0.052)
Observations	14,705	4,807	3,566	2,779	1,128	2,338
R^2	0.624	0.742	0.713	0.590	0.417	0.572
Village FEs	Y	Y	Y	Y	Y	Y
Season FEs	Y	Y	Y	Y	Y	Y
			First Stage: F statistics			
Land	77.0	62.0	40.3	37.8	15.7	19.3
Labor	49.3	34.7	17.7	25.2	12.9	14.8
Intermediates	35.8	31.7	21.5	19.9	8.9	11.8
K-Paap Wald F statistic	51.1	40.4	16.0	30.8	12.4	12.7

back

Production Functions: Pairwise Equality Test P-Values

γ

	rice	wheat	cereals	pulses
veg, frt, oil	<0.01	0.95	0.62	0.73
pulses	0.05	0.7	1	
cereals	<0.01	0.57		
wheat	<0.01			

α_{labor}

	rice	wheat	cereals	pulses
veg, frt, oil	<0.01	<0.01	0.08	0.43
pulses	0.01	0.03	0.3	
cereals	0.17	0.33		
wheat	0.65			

α_{inter}

	rice	wheat	cereals	pulses
veg, frt, oil	<0.01	<0.01	0.08	0.27
pulses	0.53	<0.01	0.94	
cereals	0.35	<0.01		
wheat	<0.01			

back

Farm: Solution

- for each farm f, need to check all 2^N possible sets of nonzero crops

back

Farm: Solution

- for each farm f, need to check all 2^N possible sets of nonzero crops
    1. fix crop set's λ_f by solving one non-linear equation for each farm

back

Farm: Solution

- for each farm f, need to check all 2^N possible sets of nonzero crops
    1. fix crop set's λ_f by solving one non-linear equation for each farm
    2. compute optimal $\{x_{f,g,i}\}_{g,i}$, $\{l_{f,i}\}_i$, $\{y_{f,i}\}_i$ given λ_f

back

Farm: Solution

- for each farm f, need to check all 2^N possible sets of nonzero crops
  1. fix crop set's λ_f by solving one non-linear equation for each farm
  2. compute optimal $\{x_{f,g,i}\}_{g,i}$, $\{l_{f,i}\}_i$, $\{y_{f,i}\}_i$ given λ_f
  3. compute optimal profit net of fixed cost
     - *expressions*

back

Farm: Solution

- for each farm f, need to check all 2^N possible sets of nonzero crops
  1. fix crop set's λ_f by solving one non-linear equation for each farm
  2. compute optimal $\{x_{f,g,i}\}_{g,i}$, $\{l_{f,i}\}_i$, $\{y_{f,i}\}_i$ given λ_f
  3. compute optimal profit net of fixed cost
     - ▶ *expressions*
- for each farm, pick profit-maximizing crop set

back

Potential Sources of Concavity

- subsistence farmer with love of variety

back

Potential Sources of Concavity

- subsistence farmer with love of variety
  - farmer minimizes costs of providing utility U:

$$U = \left(\sum_i \varphi_i y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

back

Potential Sources of Concavity

- subsistence farmer with love of variety
  - farmer minimizes costs of providing utility U:

$$U = \left(\sum_i \varphi_i y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

$\rightarrow$ mix crops to provide a diverse diet

back

Potential Sources of Concavity

- subsistence farmer with love of variety
  - farmer minimizes costs of providing utility U:

$$U = \left(\sum_i \varphi_i y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

  $\rightarrow$ mix crops to provide a diverse diet
- crop-level markets are monopolistically competitive

back

Potential Sources of Concavity

- subsistence farmer with love of variety
  - farmer minimizes costs of providing utility U:

$$U = \left(\sum_i \varphi_i y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

  $\rightarrow$ mix crops to provide a diverse diet
- crop-level markets are monopolistically competitive
  - intermediate crop aggregator combines farms' varieties:

$$Y_i = \left(\sum_f y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

back

Potential Sources of Concavity

- subsistence farmer with love of variety
  - farmer minimizes costs of providing utility U:

$$U = \left(\sum_i \varphi_i y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

$\rightarrow$ mix crops to provide a diverse diet

- crop-level markets are monopolistically competitive
  - intermediate crop aggregator combines farms' varieties:

$$Y_i = \left(\sum_f y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

$\rightarrow$ mix crops due to downward-sloping demand in each crop

back

Potential Sources of Concavity

- subsistence farmer with love of variety
  - farmer minimizes costs of providing utility U:

$$U = \left(\sum_i \varphi_i y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

  $\rightarrow$ mix crops to provide a diverse diet
- crop-level markets are monopolistically competitive
  - intermediate crop aggregator combines farms' varieties:

$$Y_i = \left(\sum_f y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

  $\rightarrow$ mix crops due to downward-sloping demand in each crop
- both setups produce farm-level FOCs that are identical to the main model

back

Potential Sources of Concavity

- subsistence farmer with love of variety
  - farmer minimizes costs of providing utility U:

$$U = \left(\sum_i \varphi_i y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

  $\rightarrow$ mix crops to provide a diverse diet
- crop-level markets are monopolistically competitive
  - intermediate crop aggregator combines farms' varieties:

$$Y_i = \left(\sum_f y_{f,i}^{\eta} \right)^{\frac{1}{\eta}}$$

  $\rightarrow$ mix crops due to downward-sloping demand in each crop
- both setups produce farm-level FOCs that are identical to the main model
  - but the appropriate way to define GE may be different

back

Splitting Distortions

- splitting $r_g \tau_{f,g} \tau_{f,g,i}$ into 3 terms is arbitrary from farm f's POV

back

Splitting Distortions

- splitting $r_g \tau_{f,g} \tau_{f,g,i}$ into 3 terms is arbitrary from farm f's POV
  - and **does not** matter for aggregate misallocation estimate

back

Splitting Distortions

- splitting $r_g \tau_{f,g} \tau_{f,g,i}$ into 3 terms is arbitrary from farm f's POV
  - and **does not** matter for aggregate misallocation estimate
  - but will matter for *partial* reallocation exercises $\rightarrow$ come up with a sensible split

back

Splitting Distortions

- splitting $r_g \tau_{f,g} \tau_{f,g,i}$ into 3 terms is arbitrary from farm f's POV
  - and **does not** matter for aggregate misallocation estimate
  - but will matter for *partial* reallocation exercises $\rightarrow$ come up with a sensible split
- split $r_g \tau_{f,g}$ from $\tau_{f,g,i}$:

back

Splitting Distortions

- splitting $r_g \tau_{f,g} \tau_{f,g,i}$ into 3 terms is arbitrary from farm f's POV
  - and **does not** matter for aggregate misallocation estimate
  - but will matter for *partial* reallocation exercises $\rightarrow$ come up with a sensible split
- split $r_g \tau_{f,g}$ from $\tau_{f,g,i}$:
  - assume $\tau_{f,g,i}$ don't distort f-level demand of g

back

Splitting Distortions

- splitting $r_g \tau_{f,g} \tau_{f,g,i}$ into 3 terms is arbitrary from farm f's POV
  - and **does not** matter for aggregate misallocation estimate
  - but will matter for *partial* reallocation exercises $\rightarrow$ come up with a sensible split
- split $r_g \tau_{f,g}$ from $\tau_{f,g,i}$:
  - assume $\tau_{f,g,i}$ don't distort f-level demand of g

  $\rightarrow$ $X_{f,g} = \sum_i x_{f,g,i} = \sum_i \tau_{f,g,i} x_{f,g,i}$

back

Splitting Distortions

- splitting $r_g \tau_{f,g} \tau_{f,g,i}$ into 3 terms is arbitrary from farm f's POV
  - and **does not** matter for aggregate misallocation estimate
  - but will matter for *partial* reallocation exercises $\rightarrow$ come up with a sensible split
- split $r_g \tau_{f,g}$ from $\tau_{f,g,i}$:
  - assume $\tau_{f,g,i}$ don't distort f-level demand of g

  $\rightarrow$ $X_{f,g} = \sum_i x_{f,g,i} = \sum_i \tau_{f,g,i} x_{f,g,i}$
- likewise restrict $\tau_{f,l,i}$:
  - assume $\tau_{f,l,i}$ don't distort f-level demand of l

back

Splitting Distortions

- splitting $r_g \tau_{f,g} \tau_{f,g,i}$ into 3 terms is arbitrary from farm f's POV
  - and **does not** matter for aggregate misallocation estimate
  - but will matter for *partial* reallocation exercises $\rightarrow$ come up with a sensible split
- split $r_g \tau_{f,g}$ from $\tau_{f,g,i}$:
  - assume $\tau_{f,g,i}$ don't distort f-level demand of g

  $\rightarrow$ $X_{f,g} = \sum_i x_{f,g,i} = \sum_i \tau_{f,g,i} x_{f,g,i}$
- likewise restrict $\tau_{f,l,i}$:
  - assume $\tau_{f,l,i}$ don't distort f-level demand of l

  $\rightarrow$ $L_f = \sum_i l_{f,i} = \sum_i \tau_{f,l,i} l_{f,i}$

back

Splitting Distortions

- splitting $r_g \tau_{f,g} \tau_{f,g,i}$ into 3 terms is arbitrary from farm f's POV
  - and **does not** matter for aggregate misallocation estimate
  - but will matter for *partial* reallocation exercises $\rightarrow$ come up with a sensible split
- split $r_g \tau_{f,g}$ from $\tau_{f,g,i}$:
  - assume $\tau_{f,g,i}$ don't distort f-level demand of g

  $\rightarrow$ $X_{f,g} = \sum_i x_{f,g,i} = \sum_i \tau_{f,g,i} x_{f,g,i}$
- likewise restrict $\tau_{f,l,i}$:
  - assume $\tau_{f,l,i}$ don't distort f-level demand of l

  $\rightarrow$ $L_f = \sum_i l_{f,i} = \sum_i \tau_{f,l,i} l_{f,i}$
- split of r_g from $\tau_{f,g}$ imposed by GE

back

Estimate the Elasticity of Substitution

- from consumption FOC:

$$\log\left(\frac{p_i C_i}{\sum_j p_j C_j}\right) = -\log\left(\sum_j \varphi_j^{\sigma} p_j^{1-\sigma}\right) + (1-\sigma)\log p_i + \sigma \log \varphi_i$$

Estimate the Elasticity of Substitution

- from consumption FOC:

$$\log\left(\frac{p_i C_i}{\sum_j p_j C_j}\right) = -\log\left(\sum_j \varphi_j^{\sigma} p_j^{1-\sigma}\right) + (1-\sigma)\log p_i + \sigma \log \varphi_i$$

- estimate with 2SLS: $\log s_{h,i} = \beta_0 + \beta_1 \log p_{h,i} + \gamma_i + \varepsilon_{h,i}$

Estimate the Elasticity of Substitution

- from consumption FOC:

$$\log\left(\frac{p_i C_i}{\sum_j p_j C_j}\right) = -\log\left(\sum_j \varphi_j^{\sigma} p_j^{1-\sigma}\right) + (1-\sigma)\log p_i + \sigma \log \varphi_i$$

- estimate with 2SLS: $\log s_{h,i} = \beta_0 + \beta_1 \log p_{h,i} + \gamma_i + \varepsilon_{h,i}$

	$\log s_{h,i}$
σ	1.699
$\log p_{h,i}$	-0.699
	(0.067)
Observations	40,833
Kleibergen-Paap F stat	230.9

Village-level instruments: Elevation × rain, ruggedness × rain, pucca roads availability

back

Choose η That Minimizes Implied Dispersion

- farm-crop revenue:

$$p_i y_{f,i} = \underbrace{\left(\left(\frac{1}{\lambda_f \tau_{f,m,i}}\right)^{\gamma_i} \Pi_g \left(\frac{1}{\tau_{f,g}\tau_{f,g,i}}\right)^{\alpha_{g,i}} \right)}_{\text{composite distortion, } dist_{f,i}}^{\frac{1}{1-\eta(\sum_g \alpha_{g,i}+\gamma_i)}} \underbrace{\left(p_i z_{f,i} \gamma_i^{\gamma_i} \eta^{\sum_g \alpha_{g,i}+\gamma_i} \Pi \left(\frac{\alpha_{g,i}}{r_g}\right)^{\alpha_{g,i}} \right)^{\frac{1}{1-\eta(\sum_g \alpha_{g,i}+\gamma_i)}}}_{\text{"objective" factors}}$$

back

Choose η That Minimizes Implied Dispersion

- farm-crop revenue:

$$p_i y_{f,i} = \left(\underbrace{\left(\frac{1}{\lambda_f \tau_{f,m,i}}\right)^{\gamma_i} \Pi_g \left(\frac{1}{\tau_{f,g}\tau_{f,g,i}}\right)^{\alpha_{g,i}}}_{\text{composite distortion, } dist_{f,i}} \right)^{\frac{1}{1-\eta(\sum_g \alpha_{g,i}+\gamma_i)}} \underbrace{\left(p_i z_{f,i} \gamma_i^{\gamma_i} \eta^{\sum_g \alpha_{g,i}+\gamma_i} \Pi \left(\frac{\alpha_{g,i}}{r_g}\right)^{\alpha_{g,i}} \right)^{\frac{1}{1-\eta(\sum_g \alpha_{g,i}+\gamma_i)}}}_{\text{“objective” factors}}$$

- $Var(\log dist_{f,i})$ needed to match observed output dispersion depends on chosen η

back

GE Profits

$$\Pi = \sum_f \left[\sum_{i=1}^{N} p_i y_{f,i} - \sum_{g=1}^{G} r_g \sum_{i=1}^{N} x_{f,g,i} \right]$$

- distortions τ, fixed costs ω, concavity η are not reflected in dividends sent to consumer
  - farmers act *as if* frictions they face had monetary representations
  - but these are non-monetary and not added/subtracted from dividends
- equivalent formulation: τs, ωs are monetary taxes/subsidies, administered by consumer
  - show up in dividends and consumer's BC as government revenue/expense
- choice is arbitrary: both formulations (or any mixture) produce identical equilibrium conditions

back

Reallocation Exercise Details

- **benchmark**: equalize land and labor distortions

back

Reallocation Exercise Details

- **benchmark**: equalize land and labor distortions
  - farms can't change crop sets, but can change crop ratios

back

Reallocation Exercise Details

- **benchmark**: equalize land and labor distortions
  - farms can't change crop sets, but can change crop ratios
  - crop sets fixed $\rightarrow$ extract all needed fundamentals from data

back

Reallocation Exercise Details

- **benchmark**: equalize land and labor distortions
  - farms can't change crop sets, but can change crop ratios
  - crop sets fixed $\rightarrow$ extract all needed fundamentals from data
  - set $\tau_{f,g}\tau_{f,g,i} = 1$, $\lambda_f = \bar{\lambda}$

back

Reallocation Exercise Details

- **benchmark**: equalize land and labor distortions
  - farms can't change crop sets, but can change crop ratios
  - crop sets fixed $\rightarrow$ extract all needed fundamentals from data
  - set $\tau_{f,g}\tau_{f,g,i} = 1$, $\lambda_f = \bar{\lambda}$
  - solve for market-clearing $\{p_i\}_i, \{r_g\}_g, \bar{\lambda}$

back

Reallocation Exercise Details

- **benchmark**: equalize land and labor distortions
  - farms can't change crop sets, but can change crop ratios
  - crop sets fixed $\rightarrow$ extract all needed fundamentals from data
  - set $\tau_{f,g}\tau_{f,g,i} = 1$, $\lambda_f = \bar{\lambda}$
  - solve for market-clearing $\{p_i\}_i, \{r_g\}_g, \bar{\lambda}$
- **least-distorted state**:

back

Reallocation Exercise Details

- **benchmark**: equalize land and labor distortions
  - farms can't change crop sets, but can change crop ratios
  - crop sets fixed $\rightarrow$ extract all needed fundamentals from data
  - set $\tau_{f,g}\tau_{f,g,i} = 1$, $\lambda_f = \bar{\lambda}$
  - solve for market-clearing $\{p_i\}_i, \{r_g\}_g, \bar{\lambda}$
- **least-distorted state**:
  - Tamil Nadu has the lowest variance of the composite distortion

back

Reallocation Exercise Details

- **benchmark**: equalize land and labor distortions
  - farms can't change crop sets, but can change crop ratios
  - crop sets fixed $\rightarrow$ extract all needed fundamentals from data
  - set $\tau_{f,g}\tau_{f,g,i} = 1$, $\lambda_f = \bar{\lambda}$
  - solve for market-clearing $\{p_i\}_i, \{r_g\}_g, \bar{\lambda}$
- **least-distorted state**:
  - Tamil Nadu has the lowest variance of the composite distortion
  - compute variances of each type of distortion in Tamil Nadu

back

Reallocation Exercise Details

- **benchmark**: equalize land and labor distortions
  - farms can't change crop sets, but can change crop ratios
  - crop sets fixed $\rightarrow$ extract all needed fundamentals from data
  - set $\tau_{f,g}\tau_{f,g,i} = 1$, $\lambda_f = \bar{\lambda}$
  - solve for market-clearing $\{p_i\}_i, \{r_g\}_g, \bar{\lambda}$
- **least-distorted state**:
  - Tamil Nadu has the lowest variance of the composite distortion
  - compute variances of each type of distortion in Tamil Nadu
  - downscale each Indian farm's frictions s.t. aggregate variances match Tamil Nadu variances

back

Benchmark Exercise Sensitivity

η \ σ	1.1	1.3	1.5	1.7	1.9	2.1	2.3
0.93	1588	1781	1961	2123	2267	2393	2503
0.83	1300	1472	1635	1787	1924	2047	2157
0.73	762	837	909	977	1040	1098	1152
0.63	277	284	289	*294*	299	303	306
0.53	139	140	140	141	141	142	142
0.43	99	99	99	99	99	100	100
0.33	79	79	80	80	80	80	80

(a) reallocation gain, %

back

Benchmark Exercise Sensitivity

η \ σ	1.1	1.3	1.5	1.7	1.9	2.1	2.3
0.93	1588	1781	1961	2123	2267	2393	2503
0.83	1300	1472	1635	1787	1924	2047	2157
0.73	762	837	909	977	1040	1098	1152
0.63	277	284	289	*294*	299	303	306
0.53	139	140	140	141	141	142	142
0.43	99	99	99	99	99	100	100
0.33	79	79	80	80	80	80	80

(a) reallocation gain, %

- misallocation estimates are always sensitive to calibrated concavity

back

Benchmark Exercise Sensitivity

η \ σ	1.1	1.3	1.5	1.7	1.9	2.1	2.3
0.93	1588	1781	1961	2123	2267	2393	2503
0.83	1300	1472	1635	1787	1924	2047	2157
0.73	762	837	909	977	1040	1098	1152
0.63	277	284	289	*294*	299	303	306
0.53	139	140	140	141	141	142	142
0.43	99	99	99	99	99	100	100
0.33	79	79	80	80	80	80	80

(a) reallocation gain, %

- misallocation estimates are always sensitive to calibrated concavity
  - firms can expand grow more easily in reallocation → greater gain

back

Benchmark Exercise Sensitivity

η \ σ	1.1	1.3	1.5	1.7	1.9	2.1	2.3
0.93	1588	1781	1961	2123	2267	2393	2503
0.83	1300	1472	1635	1787	1924	2047	2157
0.73	762	837	909	977	1040	1098	1152
0.63	277	284	289	294	299	303	306
0.53	139	140	140	141	141	142	142
0.43	99	99	99	99	99	100	100
0.33	79	79	80	80	80	80	80

(a) reallocation gain, %

η \ σ	1.1	1.3	1.5	1.7	1.9	2.1	2.3
0.93	+6	-6	-14	-21	-26	-30	-33
0.83	+3	-9	-18	-25	-30	-34	-38
0.73	-38	-43	-48	-52	-55	-57	-59
0.63	-23	-25	-27	-28	-29	-30	-31
0.53	-2	-3	-3	-4	-4	-4	-5
0.43	+2	+2	+2	+2	+2	+1	+1
0.33	+2	+2	+2	+2	+2	+2	+2

(b) single-product model error, %

- misallocation estimates are always sensitive to calibrated concavity
  - firms can expand grow more easily in reallocation → greater gain
- sign and magnitude of single-product model's error also depends on calibration

back

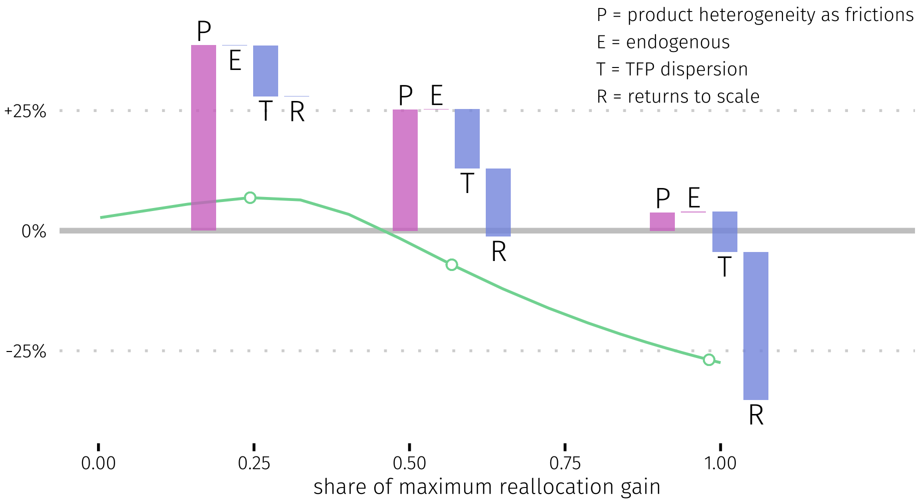
Reallocation: Role of States and Seasons

	main	within state	no split by season
multi-product:	294%	107%	314%
1-product:	212%	124%	260%

back

Remove More Distortions → 1-Product Model Overstates

- 1-product error when conducting increasingly expansive reallocations:



back